Stochastic Gene Expression in Systems Biology

Brian Munsky
Center for NonLinear Studies, Los Alamos National Lab

Mustafa Khammash
Center for Control and Dynamical Systems, University of California, Santa Barbara

On the menu...

- Today
  - Overview of Stochastic Gene Expression (Examples from the Literature)
  - Stochastic Chemical Kinetics
  - Solutions for Simple Stochastic Processes (Transcription)
  - Importance of Population Size
  - Break??
  - Moment Computations for Linear Propensities
  - Linear Noise Approximation
- Tomorrow (10:45-12:30)
  - Monte Carlo Simulation Techniques
    - ✳ Gillespie (SSA), Tau leaping, Chemical Langevin (SDEs), Slow Scale SSA.
  - Density Computations with Finite State Projection Techniques
  - Switch and Trajectory Analyses
  - Examples

Key Words

Markov Chain

Propensity Function (stochastic reaction rate)

Stoichiometry (reaction path)

Master Equation

Slides will be made available
online as soon as possible

Stochastic Gene Expression: An Overview

Why Are Stochastic Models Needed?

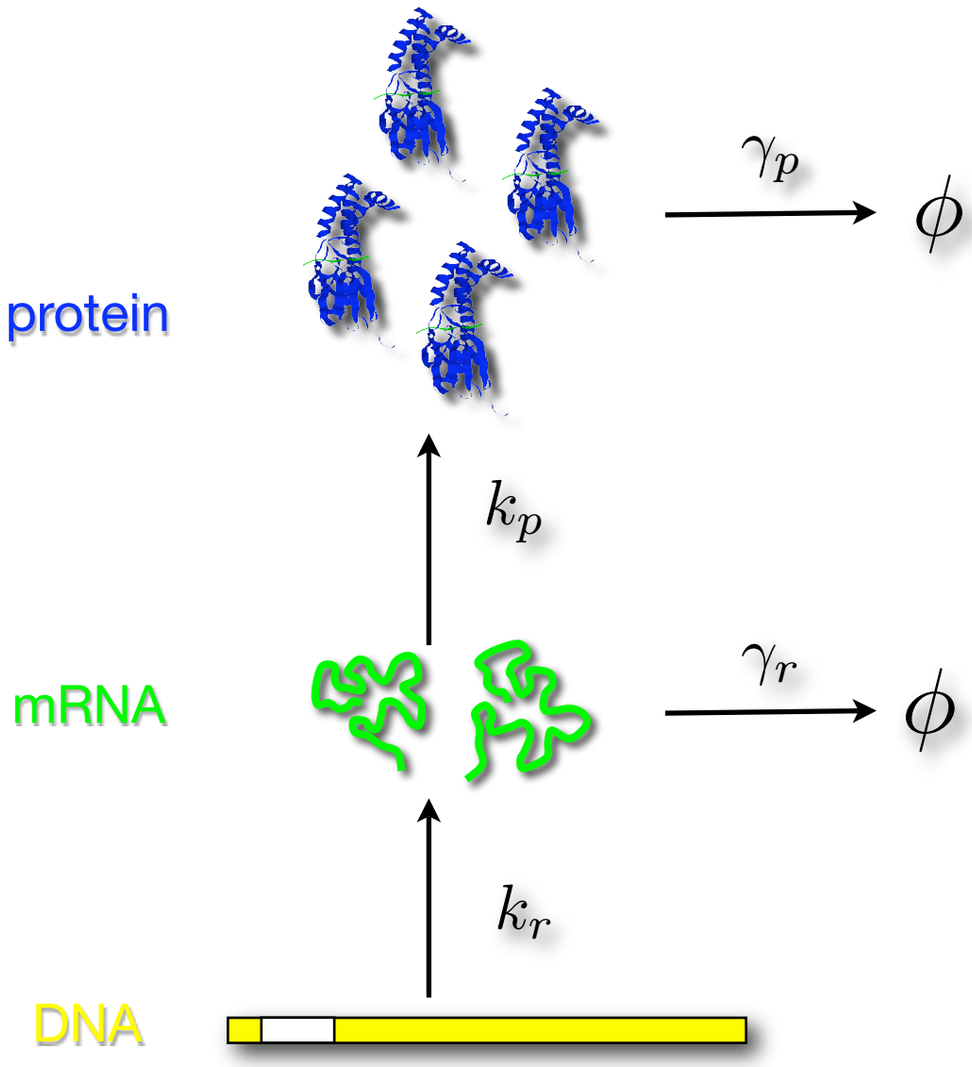
- Much of the mathematical modeling of gene networks represents gene expression deterministically

- Why worry about stochastic models?
  - Randomness
  - Quantization
  - Low copy number

- Experimental evidence indicates that stochastic fluctuations are present

- There are many examples when deterministic models are *not adequate*

The Central Dogma of Molecular Biology: Modeling Gene Expression



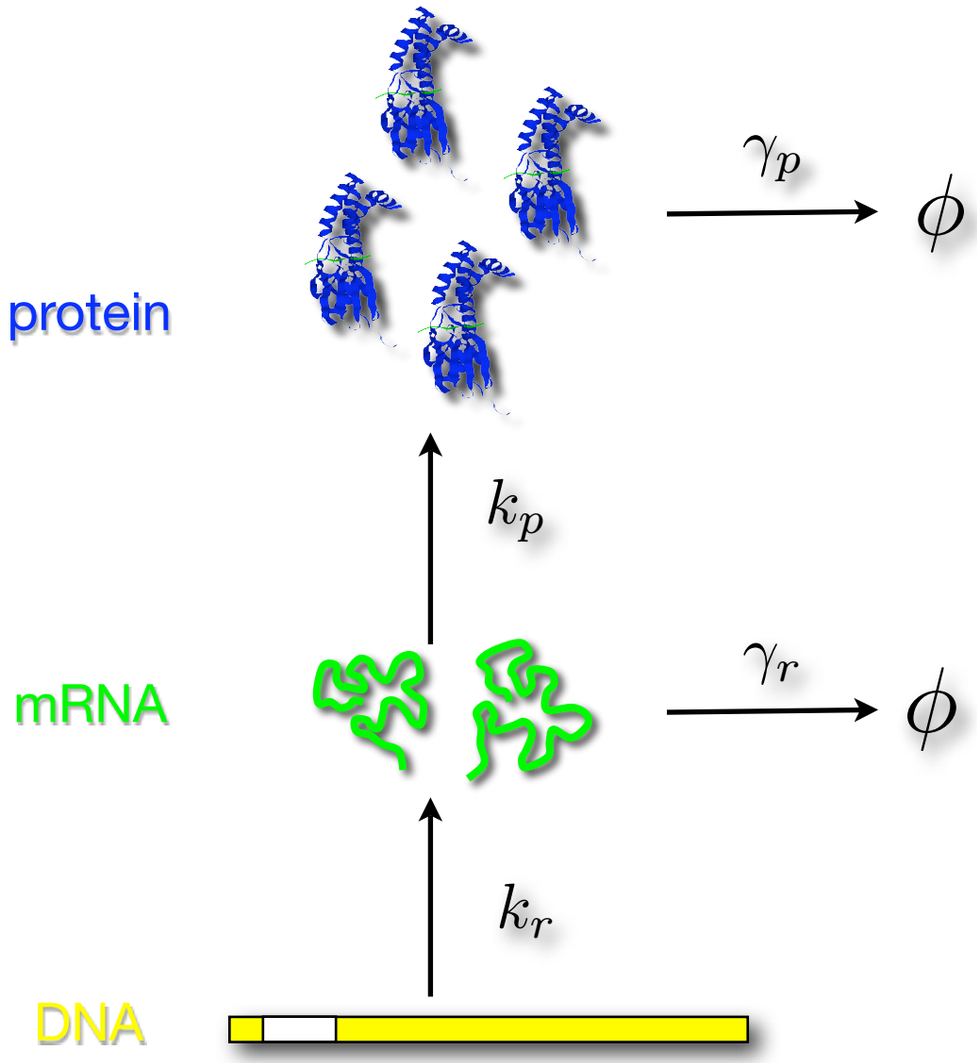
Deterministic model

$$\frac{d[mRNA]}{dt} = -\gamma_r[mRNA] + k_r$$
$$\frac{d[protein]}{dt} = -\gamma_p[protein] + k_p[mRNA]$$

Stochastic model

- Probability a single mRNA is transcribed in time dt is $k_r dt$.
- Probability a single mRNA is degraded in time dt is $(\#mRNA) \cdot \gamma_r dt$

Fluctuations at Small Copy Numbers



$$E(p) = \frac{k_r k_p}{\gamma_r \gamma_p}$$

$$C_v(p) = \frac{1}{\sqrt{E(p)}}\left(1 + \frac{kp}{\gamma_p + \gamma_r}\right)^{1/2} \quad \text{(protein)}$$

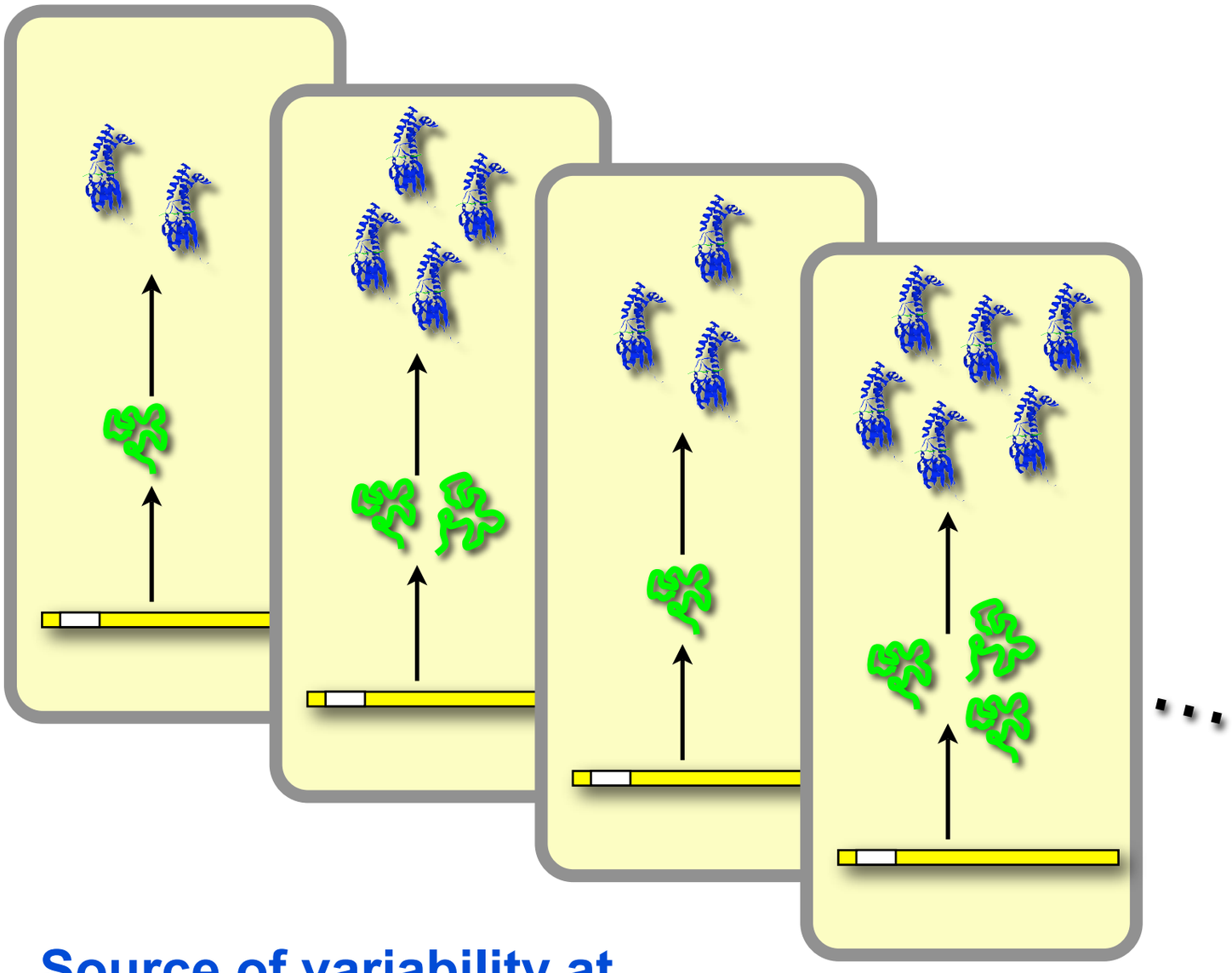
$$E(r) = \frac{k_r}{\gamma_r}$$

$$C_v(r) = \frac{1}{\sqrt{E(r)}} \quad \text{(mRNA)}$$

C_v = coefficient of variation = $\dfrac{\text{standard deviation}}{\text{mean}}$

- ★ Deterministic steady-state equals stochastic mean
- ★ Coefficient of variation goes as $1/\sqrt{\text{mean}}$
- ★ When mean is large, the coefficient of variation is (relatively) small

Intrinsic Variability in Gene Expression



Source of variability at cellular level….

- Small # of molecules
- Random events

“Intrinsic noise”

Impact of variability

- Noise propagates through the network
- Its amount depends on
  - # of molecules
  - stoichiometry
  - regulation
  - ...
- Sometimes it is suppressed; other times it is exploited
- Deterministic models are not adequate

Stochastic Influences on Phenotype

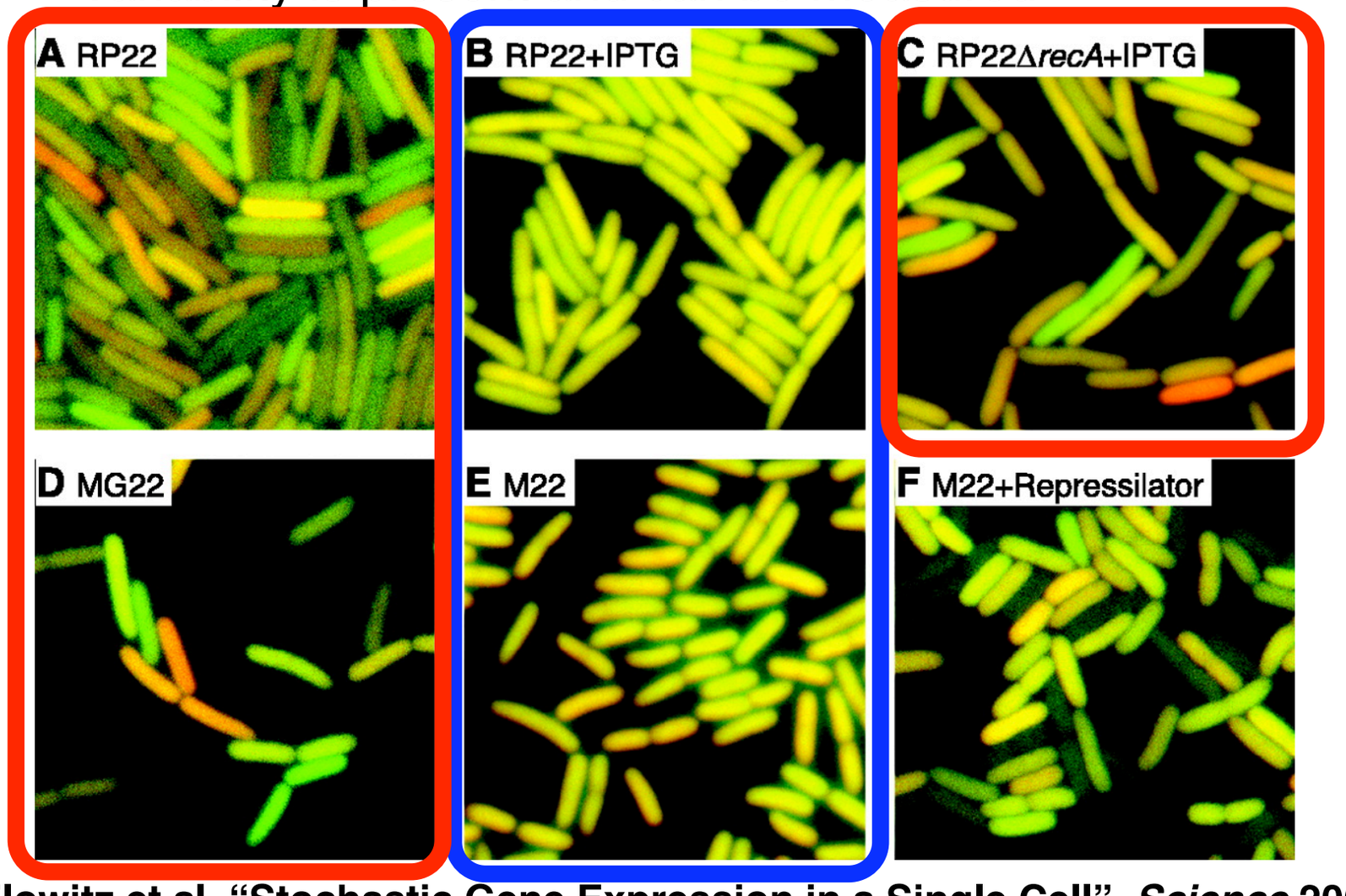
Fingerprints of identical twins

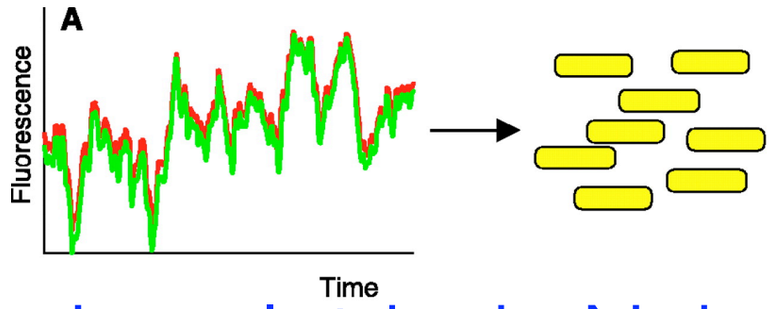
Cc, the first cloned cat and her genetic mother, Rainbow

J. Raser and E. O'Shea, "Noise in Gene Expression: Origins, Consequences, and Control", *Science*, 2005

We Are Starting to See the “Noise”!

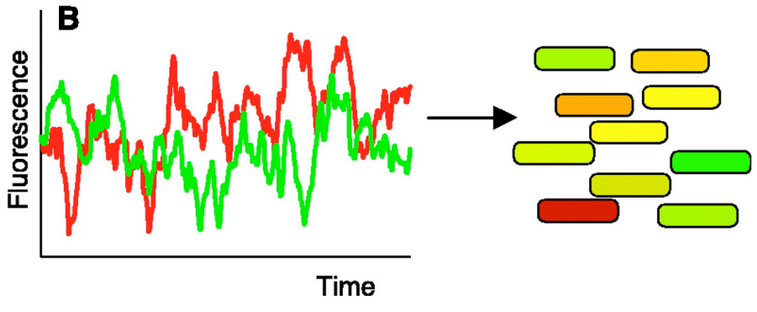
- Variability is present and can be measured





Low Intrinsic Noise

High Intrinsic Noise



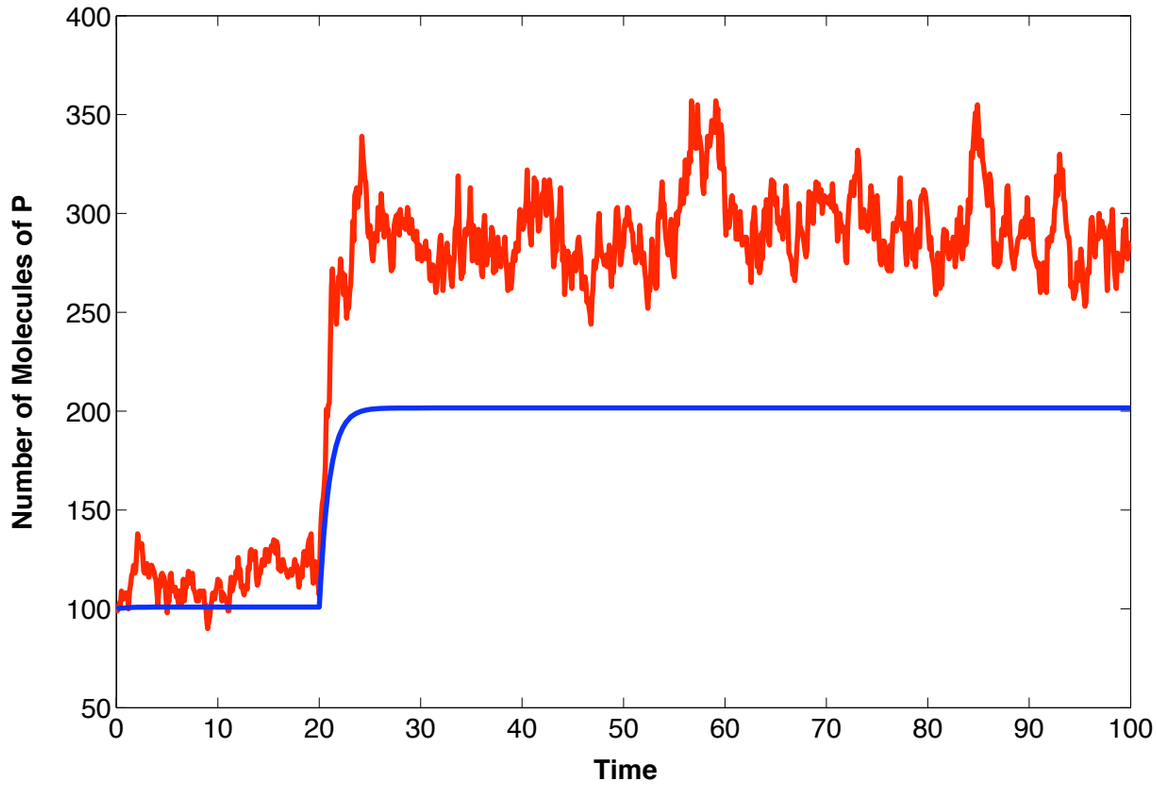
Elowitz et al, “Stochastic Gene Expression in a Single Cell”, *Science* 2002

- Inserted two reporter genes on the chromosome (cfp, yfp)
- Each was controlled by the same promoter
- Expression of cfp shown in green, yfp in red

Deterministic Model Fails to Capture Mean

$$\phi \underset{k_a S}{\overset{k}{\rightleftharpoons}} I \xrightarrow{k_p} P \xrightarrow{1} \phi$$

$$\phi \underset{k_d}{\overset{k_s}{\rightleftharpoons}} S$$

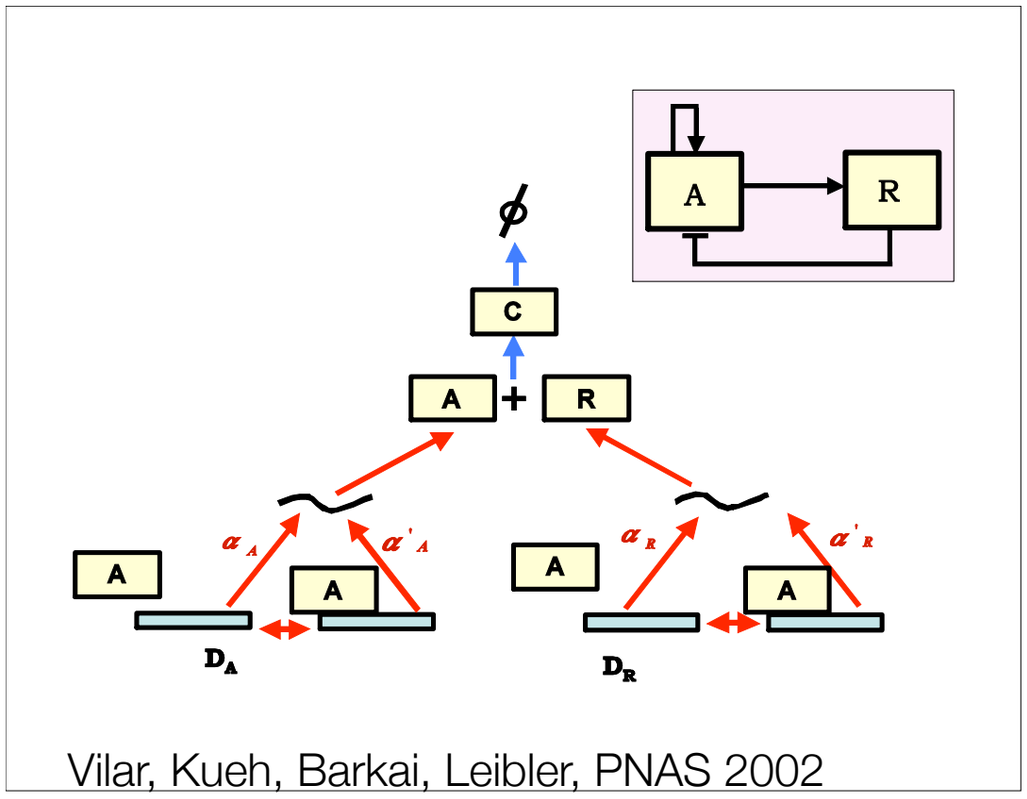


Johan Paulsson , Otto G. Berg , and Måns Ehrenberg, "Stochastic Focusing: Fluctuation-enhansed sensitivity of intracellular regulation" PNAS 2000

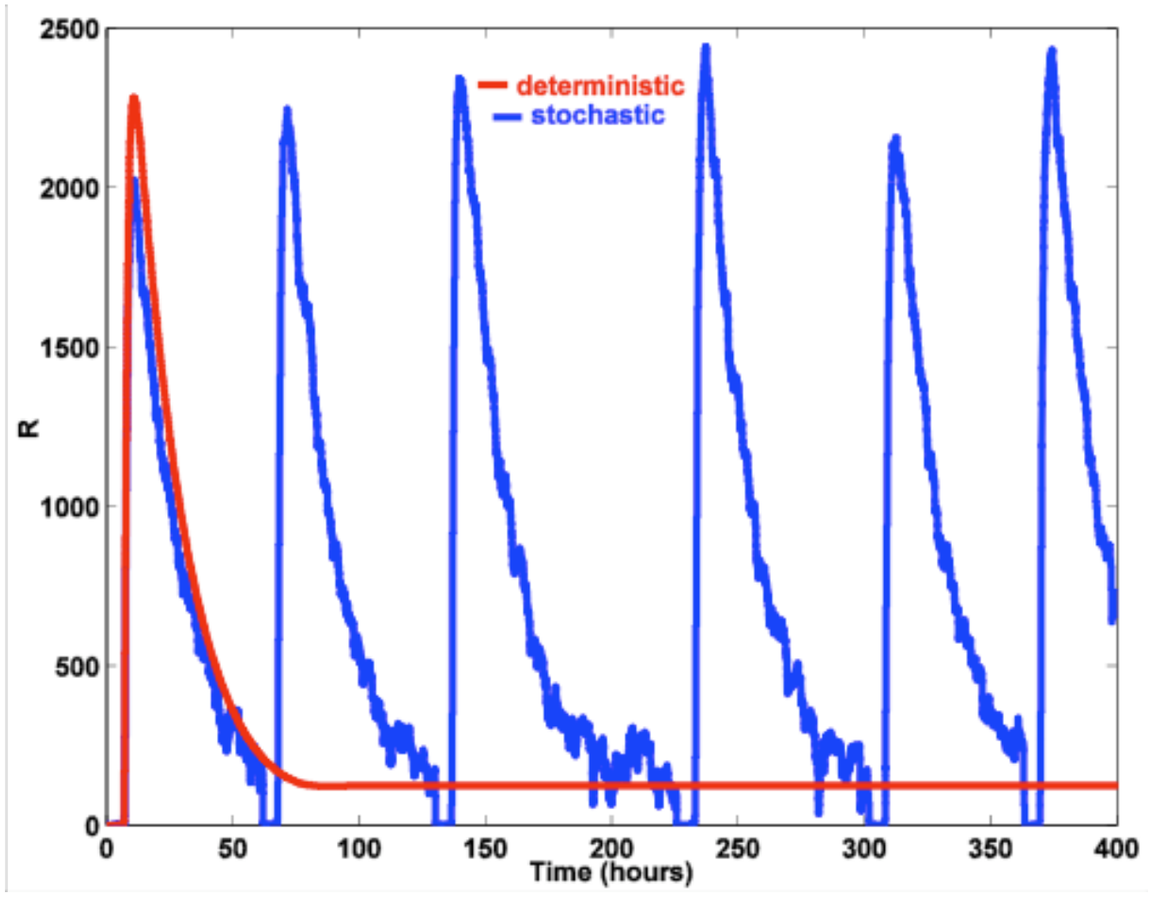
- Stochastic mean value different from deterministic steady state
- Noise *enhances* signal!

Noise Induced Oscillations

Circadian rhythm

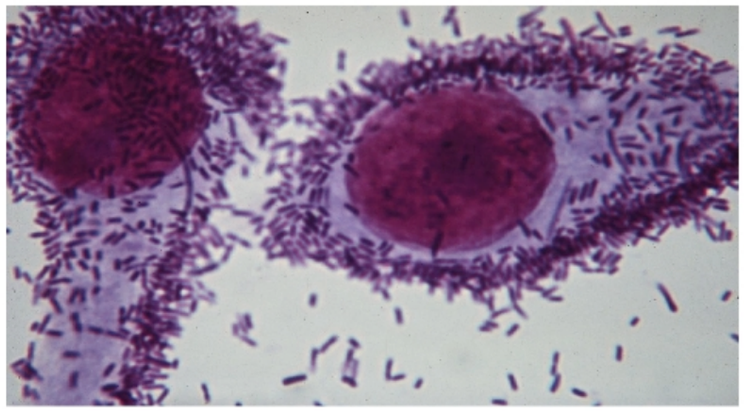


Vilar, Kueh, Barkai, Leibler, PNAS 2002



- Oscillations disappear from deterministic model after a small reduction in deg. of repressor
- (Coherence resonance) Regularity of noise induced oscillations can be manipulated by tuning the level of noise [*El-Samad, Khammash*]

The Pap Pili Stochastic Switch

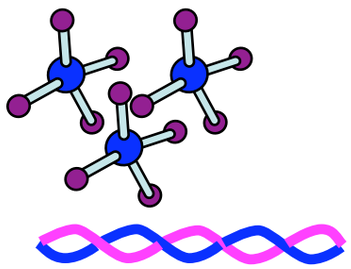


- Pili enable uropathogenic E. coli to attach to epithelial cell receptors
  - Plays an essential role in the pathogenesis of urinary tract infections
- E. coli expresses two states ON (piliated) or OFF (unpiliated)
- Piliation is controlled by a stochastic switch that involves random molecular events

The Importance of Stochasticity.

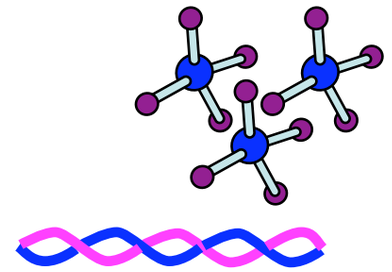
Stochastic Switching: Identical genotypes and identical environments can produce different phenotypes.

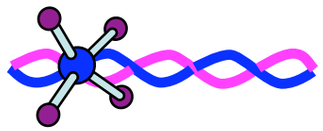
Brian Munsky

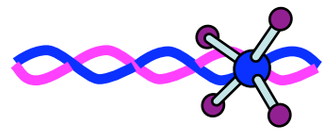


Same chemical environment.

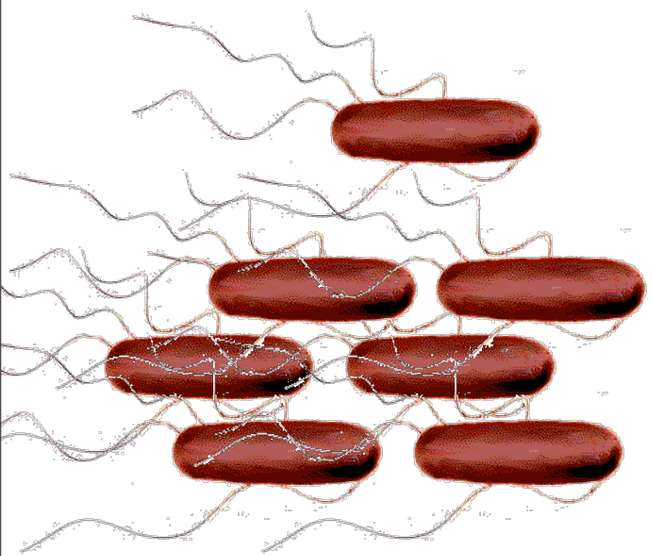
Same genetic code.



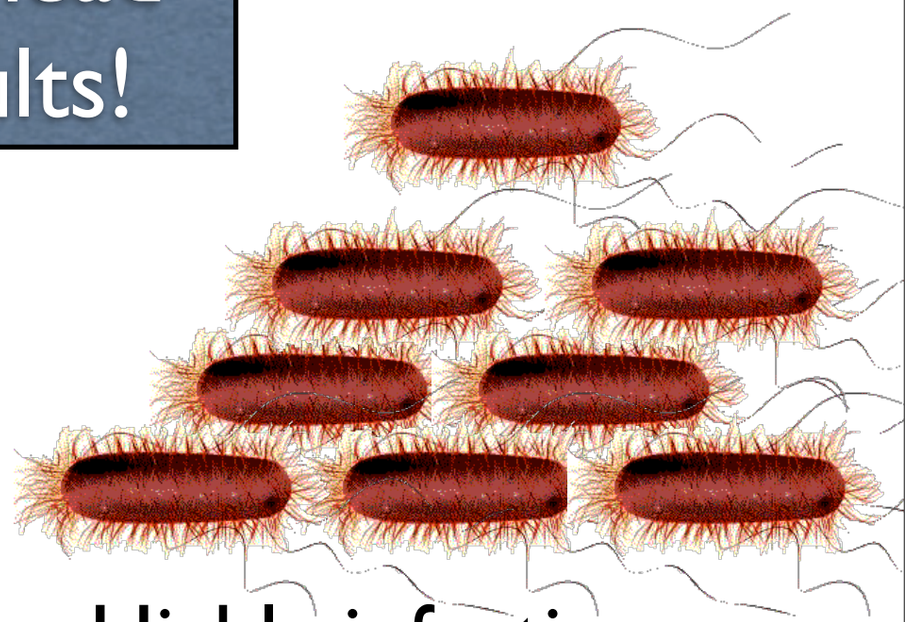




Random reactions can lead to vastly different results!



Harmless phenotype.



Highly infectious phenotype.

15

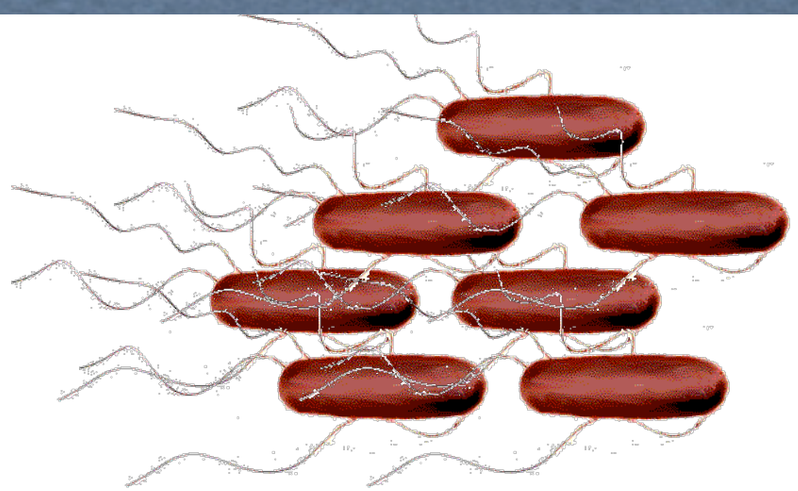
The Importance of Stochasticity.

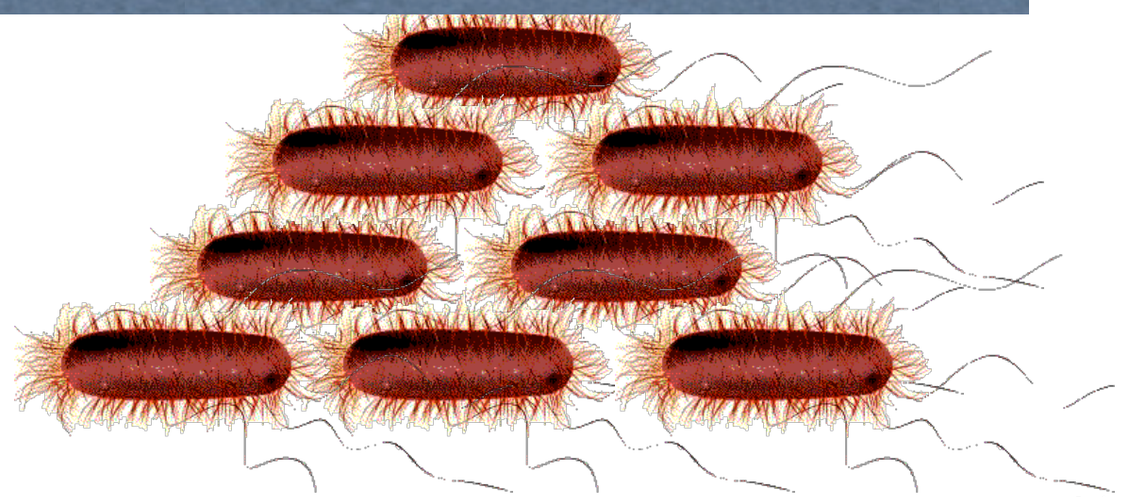
Stochastic Switching: Identical genotypes and identical environments can produce different phenotypes.

Brian Munsky

For these systems, we need analytical models to answer:

★ What will happen?
★ How frequently?
★ Why does it happen?
★ Under what conditions?
★ What advantages does it provide?
★ How can we prevent it?
★ How can we cause it?

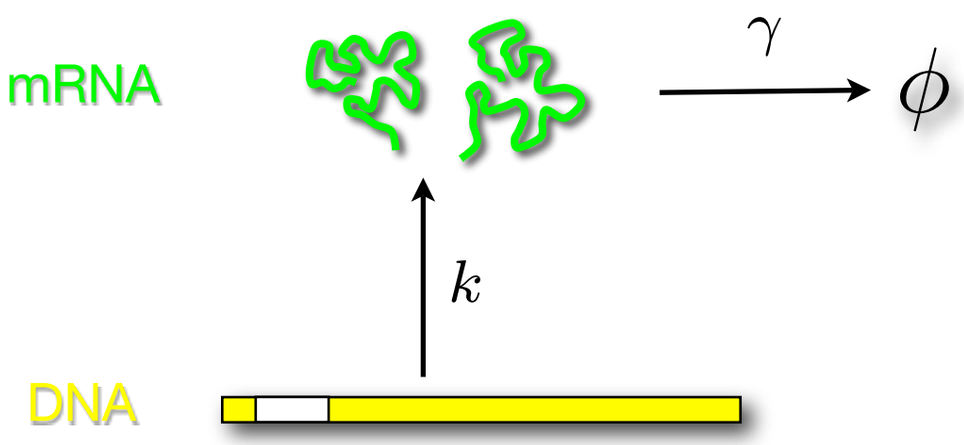




16

Markov Chain Description for Dynamical Processes

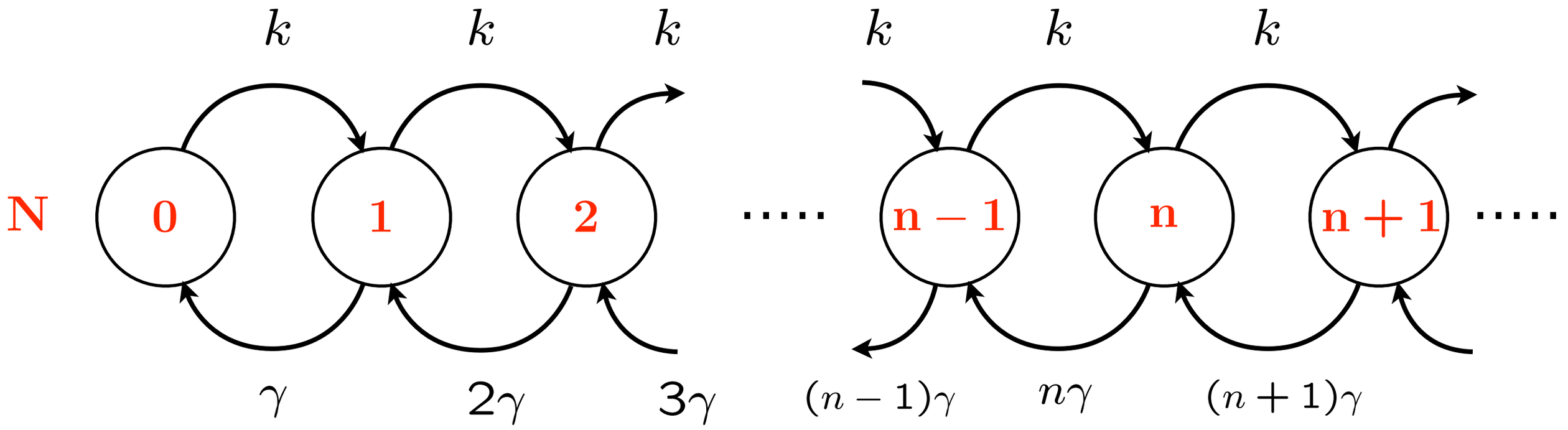
RNA Copy Number as a Random Variable



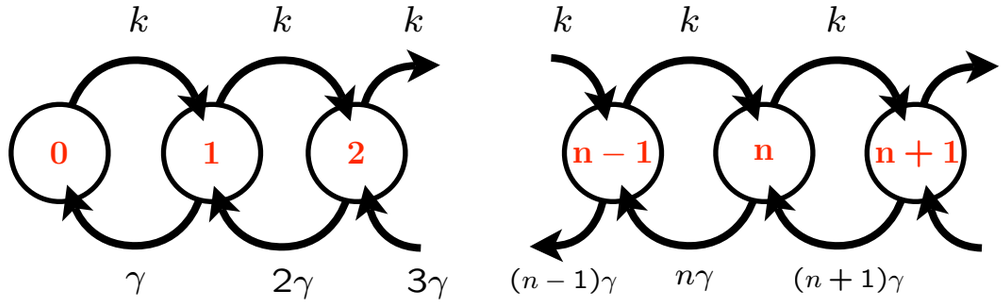
mRNA copy number $N(t)$ is a random variable

Transcription: Probability a single mRNA is transcribed in time dt is $k\ dt$

Degradation: Probability a single mRNA is degraded in time dt is $n\gamma dt$



Key Question:



Find $p(n,t)$, the probability that $N(t) = n$.

$$P(n, t+dt) = P(n-1, t) \cdot kdt$$ Prob.$\{N(t) = n-1$ and mRNA created in [t,t+dt)$\}$

$$+\ P(n+1, t) \cdot (n+1)\gamma dt$$ Prob.$\{N(t) = n+1$ and mRNA degraded in [t,t+dt)$\}$

$$+\ P(n, t) \cdot (1 - kdt)(1 - n\gamma dt)$$ Prob.$\{N(t) = n$ and mRNA not created nor degraded in [t,t+dt)$\}$

$$P(n, t+dt) - P(n, t) = P(n-1, t)kdt + P(n+1, t)(n+1)\gamma dt - P(n, t)(k + n\gamma)dt + O(dt^2)$$

Dividing by dt and taking the limit as $dt \to 0$

The Chemical Master Equation

$$\frac{d}{dt}P(n, t) = kP(n-1, t) + (n+1)\gamma P(n+1, t) - (k + n\gamma)P(n, t)$$

mRNA Stationary Distribution

We look for the stationary distribution $P(n,t) = p(n)\ \forall t$

The stationary solution satisfies: $\frac{d}{dt}P(n,t) = 0$

From the Master Equation ...

$$(k + n\gamma)p(n) = kp(n-1) + (n+1)\gamma p(n+1)$$

$n = 0$ $\quad kp(0) = \gamma p(1)$

$n = 1$ $\quad kp(1) = 2\gamma p(2)$

$n = 2$ $\quad kp(2) = 3\gamma p(3)$

$\vdots$

$$kp(n-1) = n\gamma\ p(n)$$

$kp(n-1) = n\gamma\ p(n)$ We can express $p(n)$ as a function of $p(0)$:

$$\begin{aligned} p(n) &= \frac{k}{\gamma}\frac{1}{n}\ p(n-1) \\ &= \left(\frac{k}{\gamma}\right)^2 \frac{1}{n}\ \frac{1}{n-2}\ p(n-2) \\ &\vdots \\ &= \left(\frac{k}{\gamma}\right)^n \frac{1}{n!}\ p(0) \end{aligned}$$

We can solve for $p(0)$ using the fact $\sum_{n=0}^{\infty} p(n) = 1$

$$\begin{aligned} 1 &= \sum_{n=0}^{\infty} \left(\frac{k}{\gamma}\right)^n \frac{1}{n!}\ p(0) \\ &= e^{k/\gamma}\ p(0) \quad \Longrightarrow \quad p(0) = e^{-k/\gamma} \end{aligned}$$

$$p(n) = e^{-a}\frac{a^n}{n!} \qquad a = \frac{k}{\gamma}$$

Poisson Distribution

We can compute the mean and variance of the Poisson RV $\bar{N}$ with density $p(n) = e^{-a}\frac{a^n}{n!}$:

$$\mu = E[\bar{N}] = \sum_{n=0}^{\infty} np(n) = e^{-a} \sum_{n=0}^{\infty} n\frac{a^n}{n!} = a$$

The second moment

$$E[\bar{N}^2] = \sum_{n=0}^{\infty} n^2 p(n) = a^2 + a$$

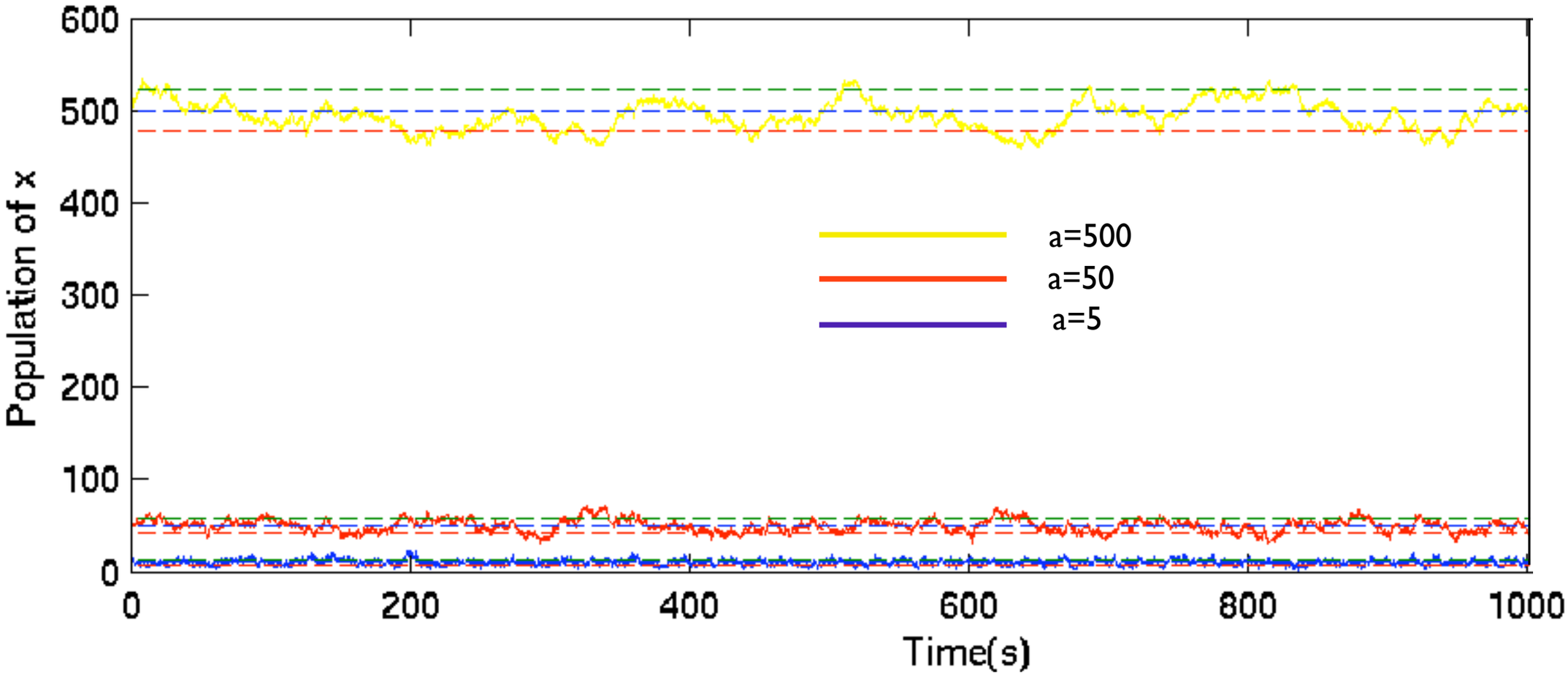
Therefore,

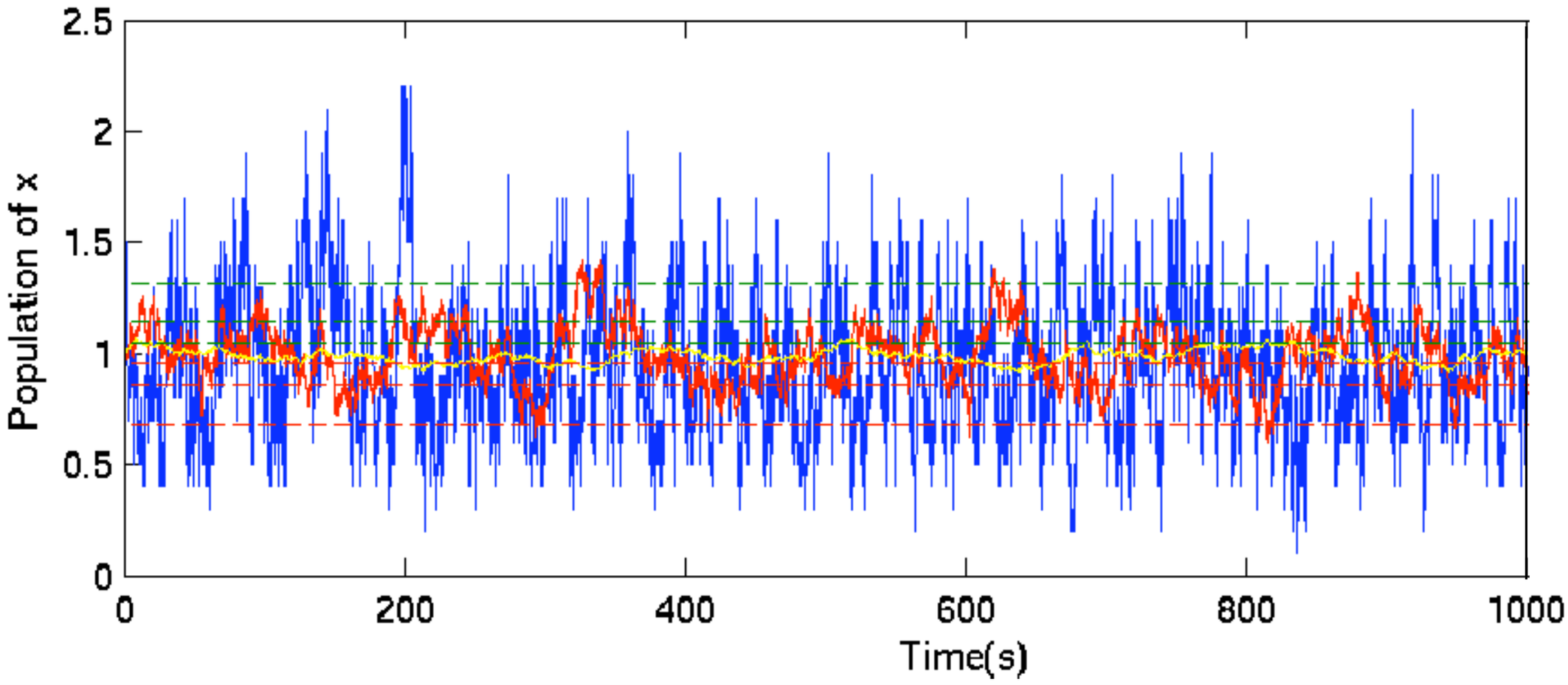
$$\sigma^2 = E[\bar{N}^2] - E[\bar{N}]^2 = a$$

mean = variance = a

The coefficient of variation $C_v = \sigma/\mu$ is

$$C_v = \frac{1}{\sqrt{a}} = \frac{1}{\sqrt{\mu}}$$


600
500
400
300
200
100
0
Population of x
0
200
400
600
800
1000
Time(s)
a=500
a=50
a=5

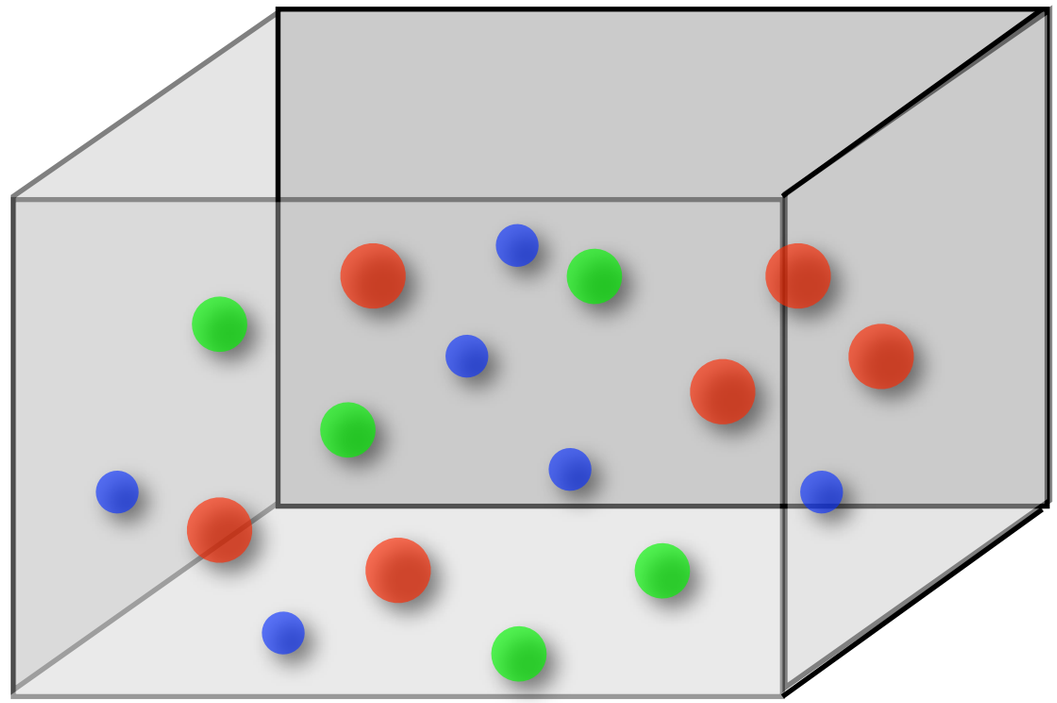

2.5
2
1.5
1
0.5
0
Population of x
0
200
400
600
800
1000
Time(s)

Stochastic Chemical Kinetics

Formulation of Stochastic Chemical Kinetics

Gillespie, Physical A, 1992

Reaction volume=Ω



Key Assumptions

(**Well-Mixed**) The probability of finding any molecule in a region $d\Omega$ is given by $\frac{d\Omega}{\Omega}$.

(**Thermal Equilibrium**) The molecules move due to the thermal energy. The reaction volume is at a constant temperature T. The velocity of a molecule is determined according to a Boltzman distribution:

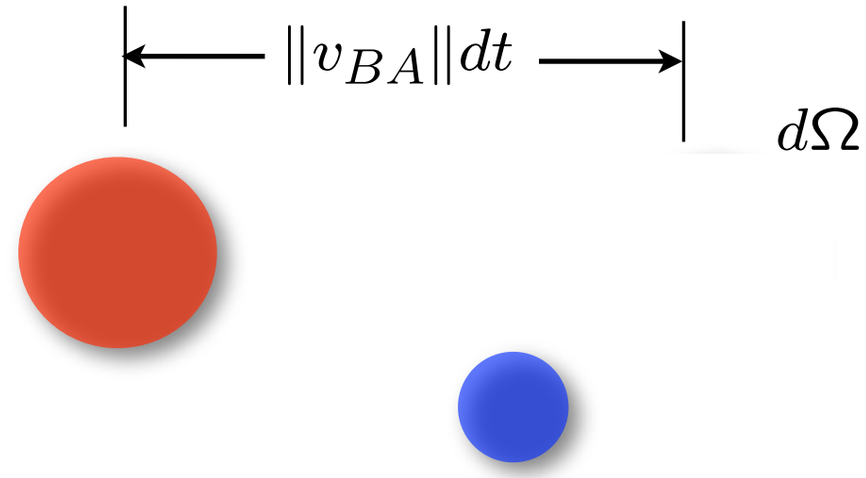
$$f_{v_x}(v) = f_{v_y}(v) = f_{v_z}(v) = \sqrt{\frac{m}{2\pi k_B T}}\, e^{-\frac{m}{2k_B T}v^2}$$

Probability of Collision: Two Specific Molecules

Given:

- Two spheres A and B with velocities v_A and v_B, and radii r_A and r_B.
- The probability that the center of either sphere lies in a volume $d\Omega$ is given by $\frac{d\Omega}{\Omega}$.

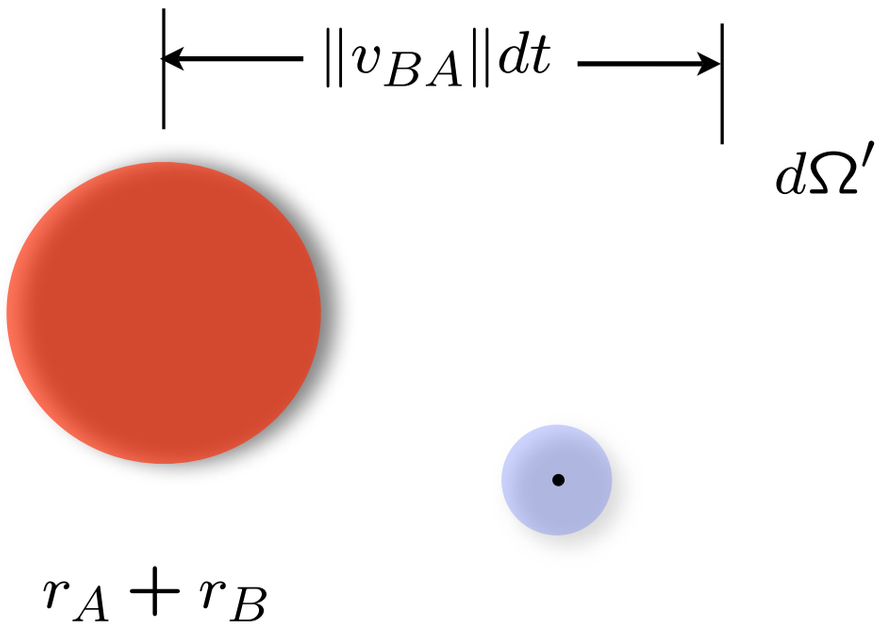
What is the probability that A and B will collide in the time $[t, t + dt]$?



In the time $[t, t + dt]$ molecule A sweeps a volume of $d\Omega = \pi r_B^2 \ \|v_{BA}\| \ dt$

Collision takes place if *any part of* A lies in the region $d\Omega$.

Equivalently . . .



During $[t, t + dt]$ a molecule with radius $r_A + r_B$ sweeps a volume of $d\Omega' = \pi(r_A + r_B)^2 \ \|v_{BA}\| \ dt$

Collision takes place if *the center of* A lies in the region $d\Omega'$.

The probability of A and B colliding during $[t, t + dt]$ is

$$\frac{1}{\Omega}\pi(r_A + r_B)^2\|v_{BA}\| \ dt$$

Note:

- The probability of A and B colliding was computed for a *given a relative velocity of* v_{BA} (conditional probability)

- The relative velocity is a *random variable*, and we must average over all velocities.

If we denote by $f_{BA}(\cdot)$ the probability density of the random variable V_{BA} we have

Collision Probability in [t,t+dt]

$$= \int_{\mathbb{R}^3} P(\text{collision in } [t, t+dt] \mid V_{BA} = v)\ f_{BA}(v)dv$$

$$= \int_{\mathbb{R}^3} \frac{1}{\Omega}\pi(r_A + r_B)^2 \|v\| dt\ f_{BA}(v)dv$$

$$= \frac{1}{\Omega}\pi(r_A + r_B)^2 dt \int_{\mathbb{R}^3} \|v\| f_{BA}(v)dv$$

mean relative speed

The probability density function of $f_{BA}(\cdot)$ can be easily computed from the Boltzman distribution of the velocity and the independence of V_x, V_y, and V_z.

$$f_{BA}(v) = \left(\frac{\widehat{m}}{2\pi k_B T}\right)^{3/2} e^{-\frac{\widehat{m}}{2k_B T}\|v\|^2}, \qquad \text{where } \widehat{m} = \frac{m_A + m_B}{2}$$

Hence

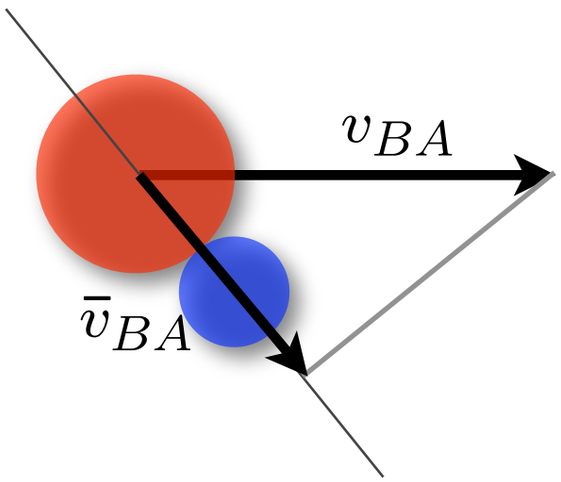
$$\begin{aligned} \text{Mean relative speed} &= \int_{\mathbb{R}^3} \|v\| f_{BA}(v) dv \\ &= \int_{\mathbb{R}^3} \|v\| \left(\frac{\widehat{m}}{2\pi k_B T}\right)^{3/2} e^{-\frac{\widehat{m}}{2k_B T}\|v\|^2} dv \\ &= \sqrt{\frac{8k_B T}{\pi \widehat{m}}} \end{aligned}$$

Probability of A-B collision within [t,t+dt]:

$$\frac{1}{\Omega}\pi(r_A + r_B)^2 dt \sqrt{\frac{8k_B T}{\pi \widehat{m}}}$$

Not all collisions lead to reactions. One can factor in the "reaction energy".

Assumption: An $A-B$ collision leads to a reaction only if the kinetic energy associated with the component of the velocity along the line of contact is greater than a critical energy ϵ.



Reaction if $\frac{1}{2}\hat{m}\bar{v}_{BA}^2 > \epsilon$

It can be shown that:

$$\text{Probability (A-B reaction} \mid \text{A-B collision)} = e^{-\frac{\epsilon}{k_B T}}$$

Probability of A-B reaction within [t,t+dt]:

$$\frac{1}{\Omega}\pi(r_A + r_B)^2 \sqrt{\frac{8k_BT}{\pi\hat{m}}} e^{-\frac{\epsilon}{k_BT}} dt$$

Given N species: $\mathcal{S}_1, \ldots, \mathcal{S}_N$ with populations $x_1, \ldots, x_N$ at time t.

Consider the bimolecular reaction channel (with distinct species):

$$R: \ \mathcal{S}_i + \mathcal{S}_j \rightarrow \text{products}$$

The number of distinct $\mathcal{S}_i - \mathcal{S}_j$ pairs that can react is: $x_i \cdot x_j$. Therefore,

Probability of an R reaction within [t,t+dt]:

$$x_i x_j \frac{1}{\Omega} \pi (r_i + r_j)^2 \sqrt{\frac{8 k_B T}{\pi \hat{m}}} e^{-\frac{\epsilon}{k_B T}} dt = w(x)$$

$w(\cdot)$ is called the **propensity function**.

Consider the bimolecular reaction channel (with same species):

$$R': \ \mathcal{S}_i + \mathcal{S}_i \rightarrow \text{products}$$

The number of distinct $\mathcal{S}_i - \mathcal{S}_i$ pairs that can react is: $\frac{x_i(x_i-1)}{2}$. Therefore,

Probability of an R' reaction within [t,t+dt]:

$$\frac{x_i(x_i-1)}{2} \ \frac{1}{\Omega} \pi r_i^2 \ \sqrt{\frac{8 k_B T}{\pi \hat{m}}} e^{-\frac{\epsilon}{k_B T}} \ dt = w(x) \ dt$$

Reactions and Propensity Functions

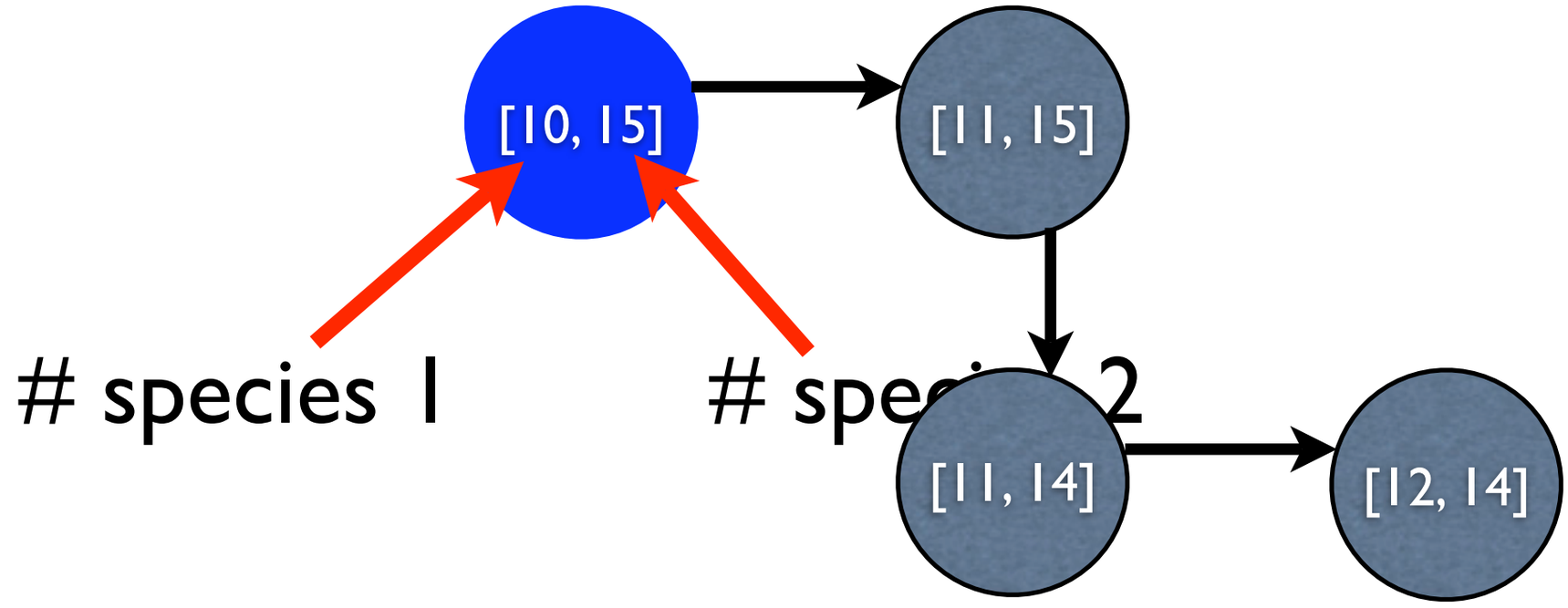
Reaction	Propensity $w(x)$	Rate c
$\phi \xrightarrow{c}$ Products	c	
$\mathcal{S}_i \xrightarrow{c}$ Products	$c \cdot x_i$	
$\mathcal{S}_i + \mathcal{S}_j \xrightarrow{c}$ Products	$c \cdot x_i x_j$	$\frac{1}{\Omega}\pi(r_i + r_j)^2 \sqrt{\frac{8k_BT}{\pi\hat{m}}} e^{-\frac{\epsilon}{k_BT}}$
$\mathcal{S}_i + \mathcal{S}_i \xrightarrow{c}$ Products	$c \cdot \frac{x_i(x_i - 1)}{2}$	$\frac{4}{\Omega}\pi r_i^2 \sqrt{\frac{8k_BT}{\pi\hat{m}}} e^{-\frac{\epsilon}{k_BT}}$

For a monomolecular reaction: c is numerically equal to the reaction rate constant k of conventional deterministic chemical kinetics

For a bimolecular reaction: c is numerically equal to k/Ω, where k is the reaction rate constant of conventional deterministic chemical kinetics

A Jump-Markov description of chemical kinetics

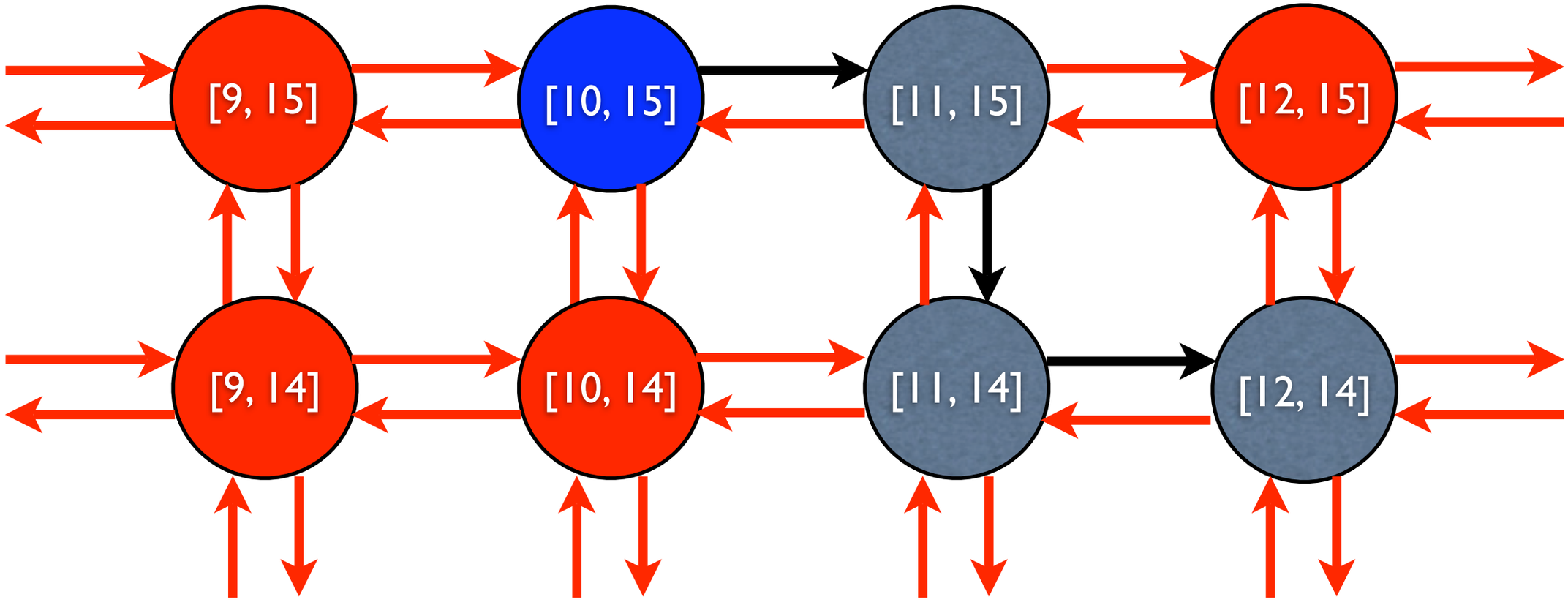
- At any time, the state of the system is defined by its integer population vector: $\mathbf{x} \in \mathbb{Z}^N$
- Reactions are transitions from one state to another:



33

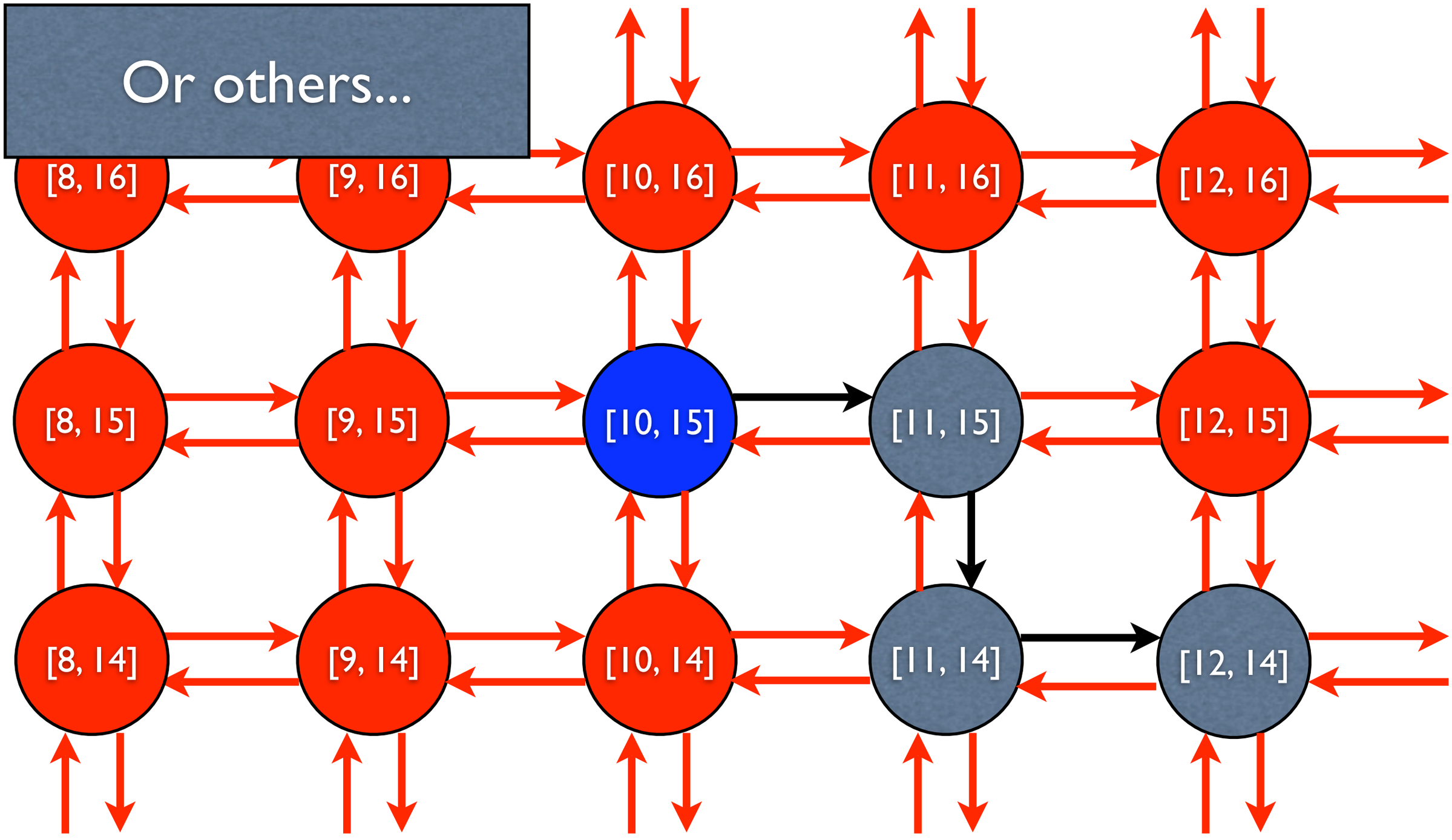
A Jump-Markov description of chemical kinetics

- At any time, the state of the system is defined by its integer population vector: $\mathbf{x} \in \mathbb{Z}^N$
- Reactions are transitions from one state to another:
- These reactions are random, others could have occurred:



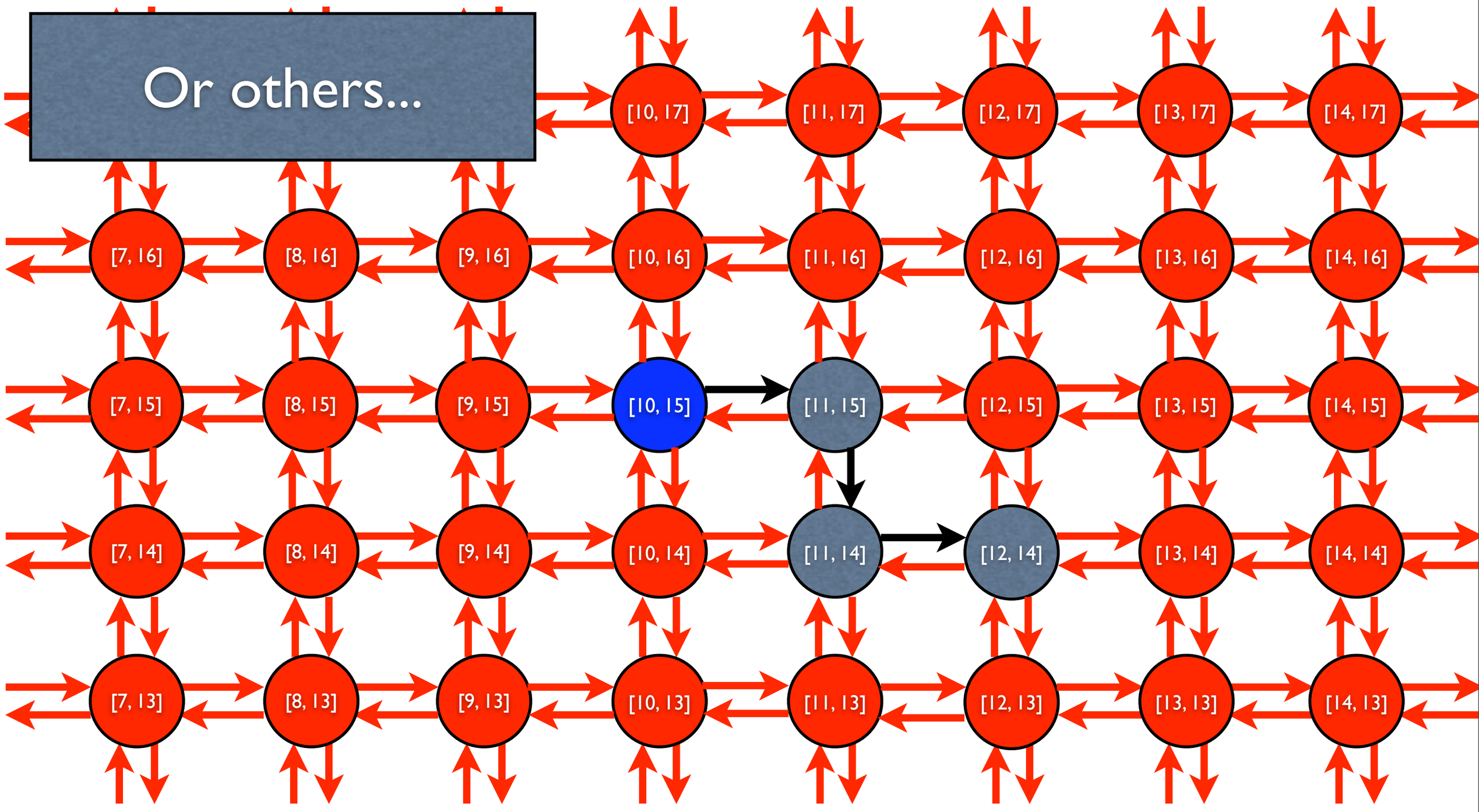
34

A Jump-Markov description of chemical kinetics



35

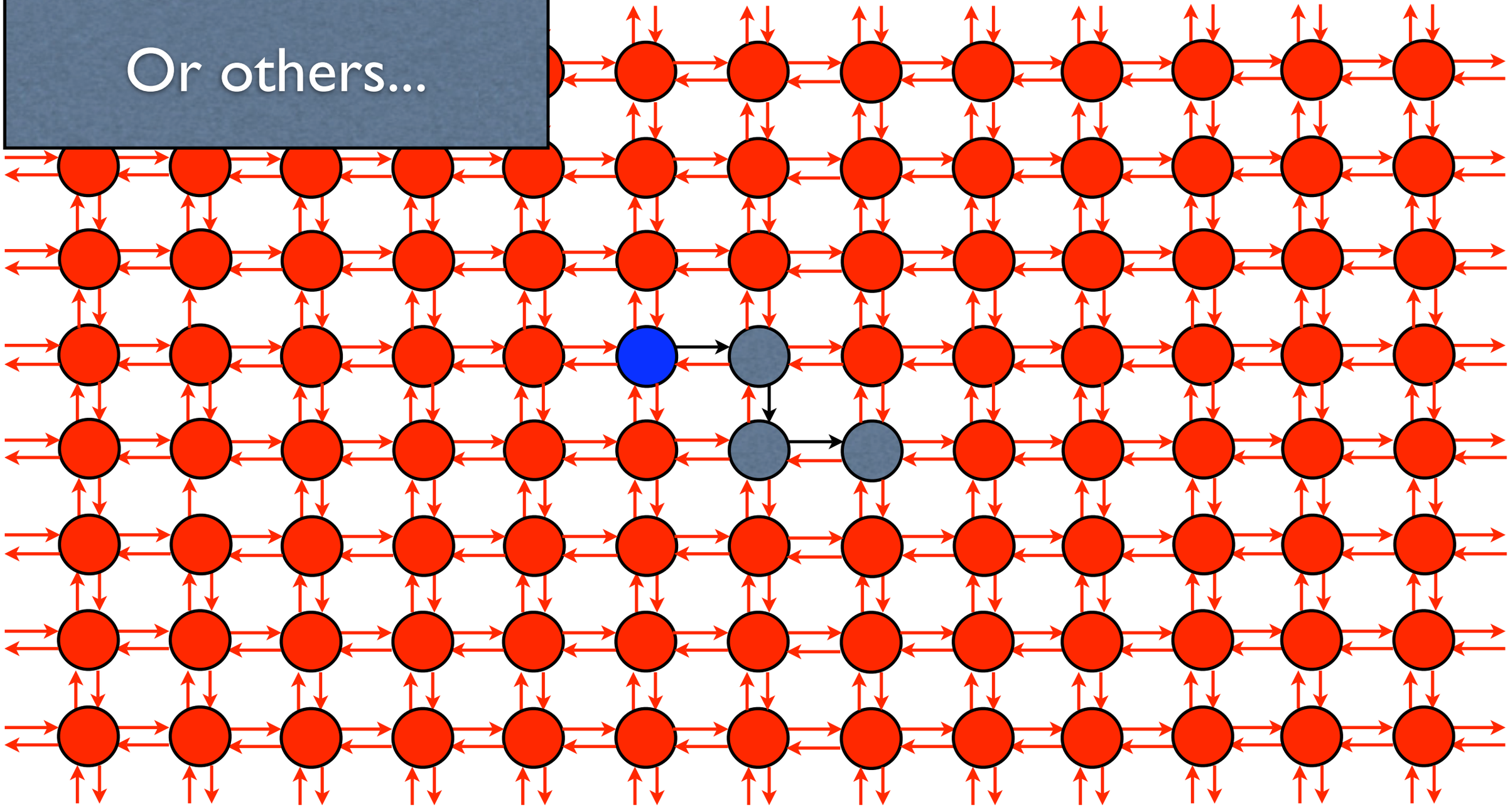
A Jump-Markov description of chemical kinetics



36

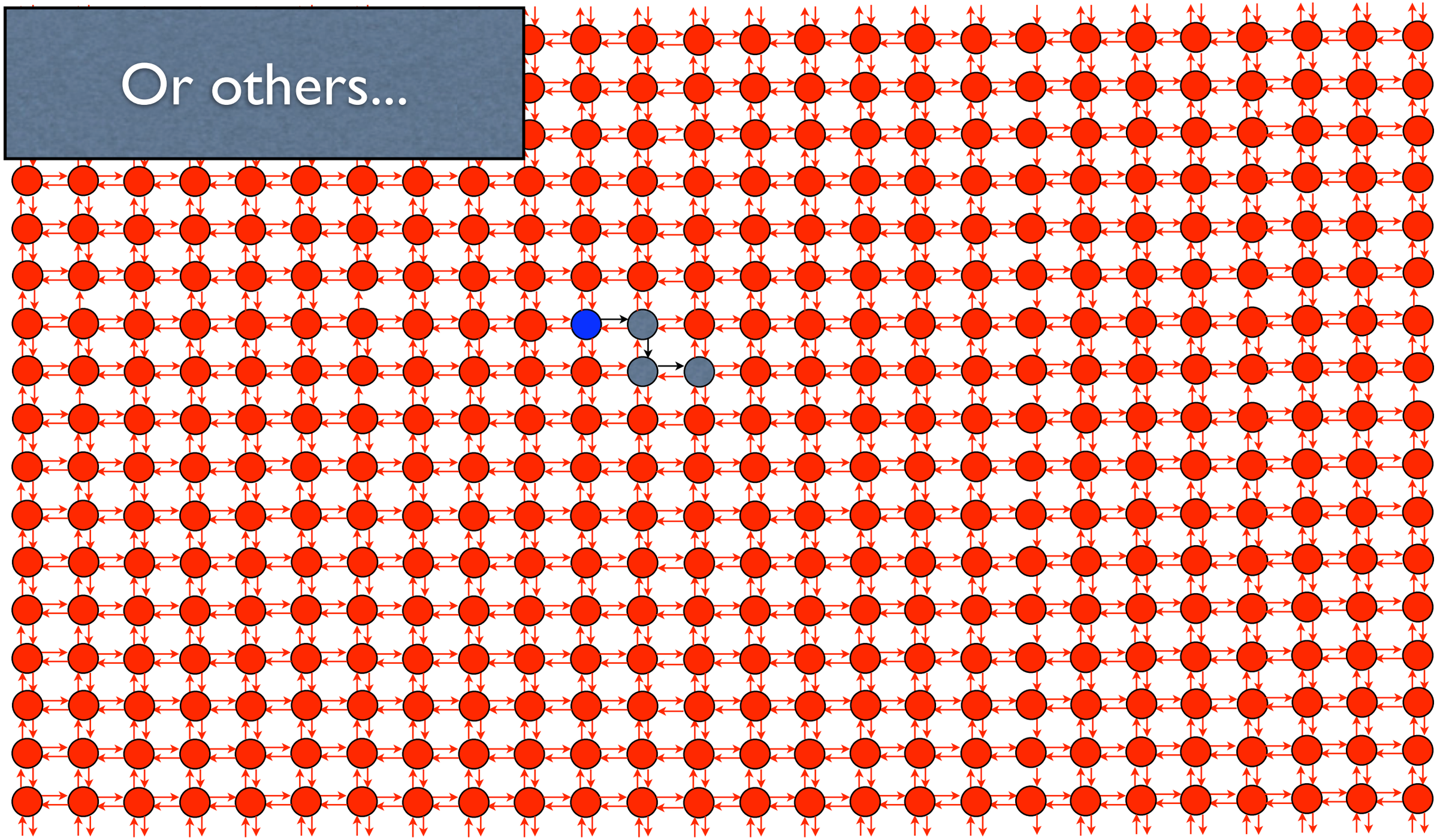
A Jump-Markov description of chemical kinetics

Or others...



37

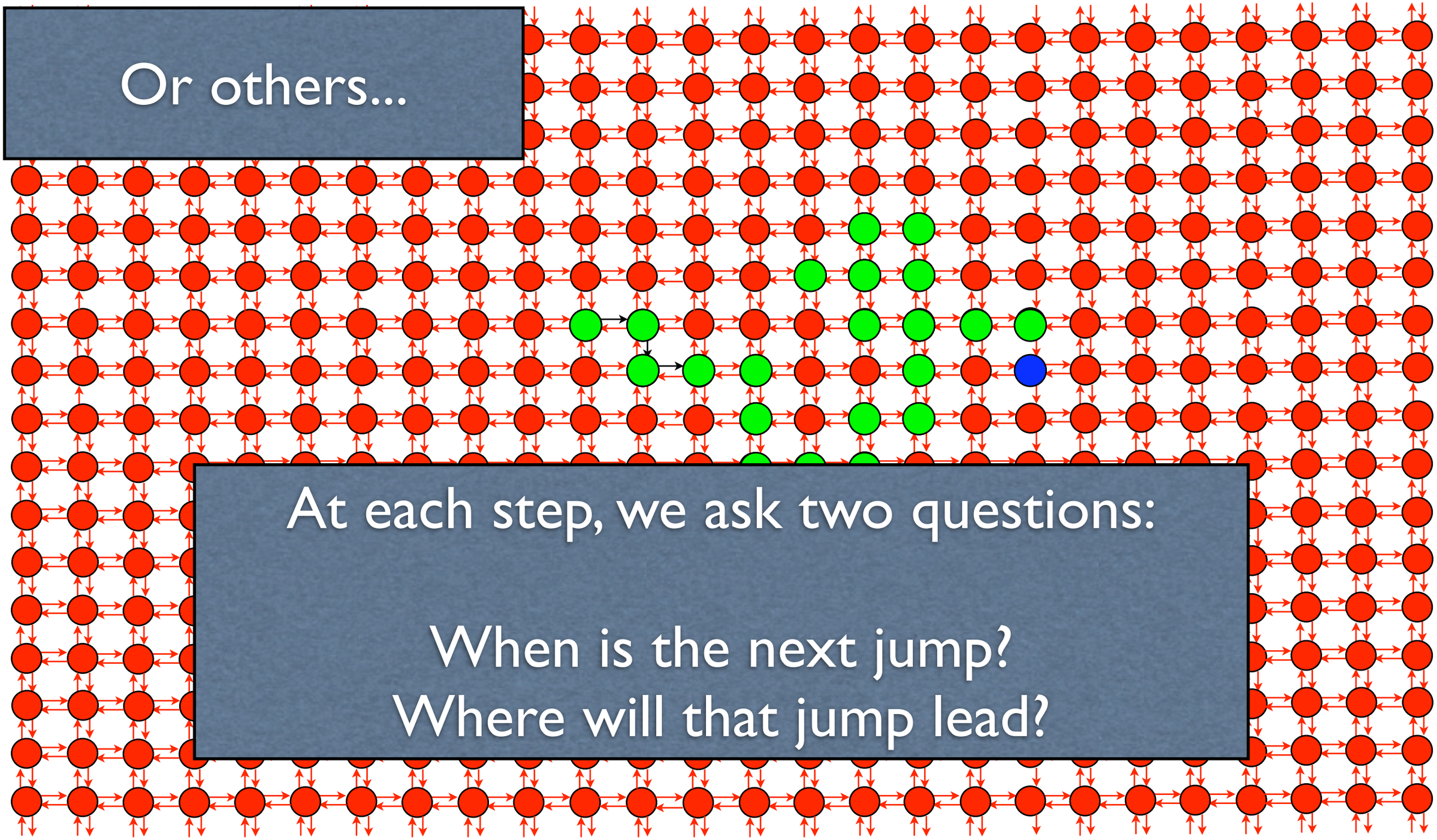
A Jump-Markov description of chemical kinetics



38

A Jump-Markov description of chemical kinetics

Or others...



At each step, we ask two questions:

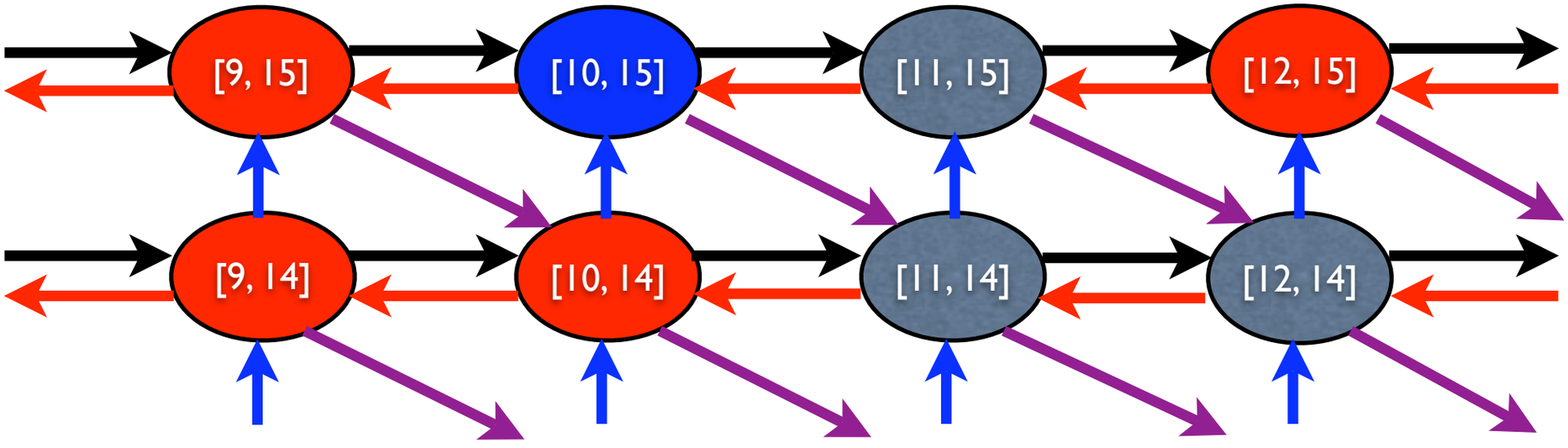
When is the next jump?
Where will that jump lead?

39

Reaction Stoichiometry (review)

- The Stoichiometric vector, **s**, refers to the relative change in the population vector after a reaction.
- There may be many different reactions for a given stoichiometry.

$\mathbf{s}_1 = [1, 0]^T$	$\mathbf{s}_2 = [-1, 0]^T$	$\mathbf{s}_3 = [0, 1]^T$	$\mathbf{s}_4 = [1, -1]^T$
$\mathcal{S}_1 \rightarrow \mathcal{S}_1 + \mathcal{S}_1$	$\mathcal{S}_1 + \mathcal{S}_1 \rightarrow \mathcal{S}_1$	$\mathcal{S}_2 \rightarrow \mathcal{S}_2 + \mathcal{S}_2$	$\mathcal{S}_2 \rightarrow \mathcal{S}_1$
$\mathcal{S}_2 \rightarrow \mathcal{S}_2 + \mathcal{S}_1$	$\mathcal{S}_1 + \mathcal{S}_2 \rightarrow \mathcal{S}_2$	$\mathcal{S}_1 \rightarrow \mathcal{S}_1 + \mathcal{S}_2$	$\mathcal{S}_1 + \mathcal{S}_2 \rightarrow \mathcal{S}_1 + \mathcal{S}_1$
$\emptyset \rightarrow \mathcal{S}_1$	$\mathcal{S}_1 \rightarrow \emptyset$	$\emptyset \rightarrow \mathcal{S}_2$	$\mathcal{S}_2 + \mathcal{S}_2 \rightarrow \mathcal{S}_1 + \mathcal{S}_2$

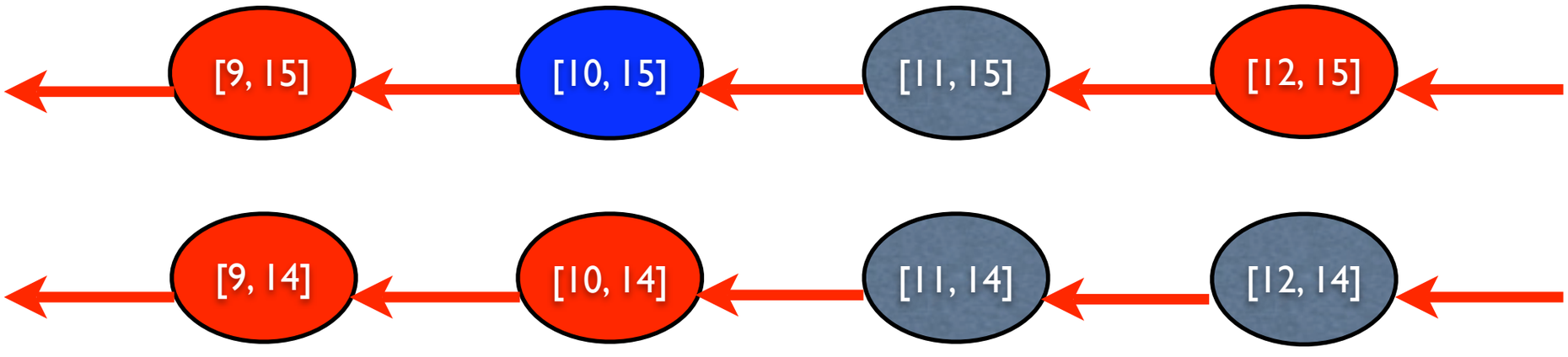


40

Reaction Propensities (review)

- The propensity, w, of a reaction is its rate.
- $\mathrm{w}_\mu dt$ is the probability that the μ^{th} reaction will occur in a time step of length dt.
- Typically, propensities depend only upon reactant populations.

$\mathbf{s}_2 = [-1, 0]^T$	$w_2(x_1, x_2)$
$\mathcal{S}_1 + \mathcal{S}_1 \rightarrow \mathcal{S}_1$	$k_1 x_2 (x_1 - 1)/2$
$\mathcal{S}_1 + \mathcal{S}_2 \rightarrow \mathcal{S}_2$	$k_2 x_1 x_2$
$\mathcal{S}_1 \rightarrow \emptyset$	$k_3 x_1$



41

The Chemical Master Equation

Prob. that no reactions fire in $[t, t+dt] = 1 - \sum_k w_k(x)dt + \mathcal{O}(dt^2)$

Prob. that reaction R_k fires once in $[t, t+dt] = w_k(x)dt + \mathcal{O}(dt^2)$

Prob. that more than one reaction fires in $[t, t+dt] = \mathcal{O}(dt^2)$

$$p(x, t+dt) = \underbrace{p(x,t)}_{\text{at } x}\underbrace{\left(1 - \sum_k w_k(x)dt + \mathcal{O}(dt^2)\right)}_{\text{No reaction fires}} + \sum_k \underbrace{p(x - s_k, t)}_{\substack{R_k \text{ reaction} \\ \text{away from } x}}\underbrace{\left(\sum_k w_k(x)dt + \mathcal{O}(dt^2)\right)}_{R_k \text{ fires once}} + \underbrace{\mathcal{O}(dt^2)}_{\substack{\text{more than one} \\ \text{reaction in } dt}}$$

$$p(x, t+dt) - p(x,t) = -p(x,t)\sum_k w_k(x)dt + \sum_k p(x - s_k, t)w_k(x)dt + \mathcal{O}(dt^2)$$

The Chemical Master Equation

$$\frac{dp(x,t)}{dt} = -p(x,t)\sum_k w_k(x) + \sum_k p(x - s_k, t)w_k(x)$$

Relationship of Stochastic and Deterministic Descriptions

Given N species $\mathcal{S}_1, \ldots, \mathcal{S}_N$ and M elementary reactions. Let $\Phi_i := [\mathcal{S}_i]$.

A deterministic description can be obtained from mass-action kinetics:

$$\frac{d\Phi}{dt} = Sf(\Phi)$$

where $f(\cdot)$ is at most a second order monomial. It depends on the type of reactions and their rates.

Example:

$$A + B \xrightarrow{k_1} C$$
$$A \xrightarrow{k_2} B$$

$$\begin{aligned}
\frac{d\Phi_A}{dt} &= -k_1\Phi_A\Phi_B - k_2\Phi_A \\
\frac{d\Phi_A}{dt} &= -k_1\Phi_A\Phi_B + k_2\Phi_A \\
\frac{d\Phi_A}{dt} &= k_1\Phi_A\Phi_B
\end{aligned}$$

or

$\frac{d\Phi}{dt} = Sf(\Phi)$ where

$$S = \begin{bmatrix} -1 & -1 \\ -1 & 1 \\ 1 & 0 \end{bmatrix}, \quad f(\Phi) = \begin{bmatrix} k_1\Phi_A\Phi_B \\ k_2\Phi_A \end{bmatrix}$$

Relationship of Stochastic and Deterministic Descriptions

Define $X^{\Omega}(t) = \frac{X(t)}{\Omega}$.

Question: How does $X^{\Omega}(t)$ relate to $\Phi(t)$?

Fact: Let $\Phi(t)$ be the deterministic solution to the reaction rate equations

$$\frac{d\Phi}{dt} = Sf(\Phi), \;\; \Phi(0) = \Phi_0.$$

Let $X^{\Omega}(t)$ be the stochastic representation of the same chemical systems with $X^{\Omega}(0) = \Phi_0$. Then for every $t \geq 0$:

$$\lim_{\Omega \to \infty} \sup_{s \leq t} \left| X^{\Omega}(s) - \Phi(s) \right| = 0 \;\; a.s.$$

Moment Computations

- Affine Propensity
- Linear Noise Approximation

Moment Computations

For the first moment $E[X_i]$, multiply the CME by x_i and sum over all $(x_1, \ldots, x_N) \in \mathbb{N}^N$

For the second moment $E[X_i X_j]$, multiply the CME by $x_i x_j$ and sum over all $(x_1, \ldots, x_N) \in \mathbb{N}^N$

$$\frac{dE[X_i]}{dt} = \sum_{k=1}^{M} s_{ik} E[w_k(X)]$$

$$\frac{dE[X_i X_j]}{dt} = \sum_{k=1}^{M} (s_{ik} E[X_j w_k(X)] + E[X_i w_k(X)] s_{jk} + s_{ik} s_{jk} E[w_k(X)])$$

Let $w(x) = [w_1(x), \ldots, w_M(x)]^T$

In matrix notation:

$$\frac{dE[X]}{dt} = SE[w(X)]$$

$$\frac{dE[XX^T]}{dt} = SE[w(X)X^T] + E[w(X)X^T]^T S^T + S\{diagE[w(X)]\}S^T$$

Affine Propensity

Suppose the propensity function is affine:

$$w(x) = Wx + w_0, \qquad (W \text{ is } N \times N,\ w_0 \text{ is } N \times 1)$$

Then $E[w(X)] = WE[X] + w_0$, and $E[w(X)X^T] = WE[XX^T] + w_0 E[X^T]$.

This gives us the moment equations:

$$\begin{aligned}
\frac{d}{dt}E[X] &= SWE[X] + Sw_0 && \text{First Moment} \\
\frac{d}{dt}E[XX^T] &= SWE[XX^T] + E[XX^T]W^T S^T + S\ diag(WE[X] + w_0)S^T \\
&+ Sw_0E[X^T] + E[X]w_0^T S^T && \text{Second Moment}
\end{aligned}$$

These are linear ordinary differential equations and can be easily solved!

Affine Propensity (cont.)

Define the covariance matrix $\Sigma = E[(X - E[X])(X - E(X)]^T]$.
We can also compute covariance equations:

$$\frac{d}{dt}\Sigma = SW\Sigma + \Sigma W^T S^T + S\ diag(WE[X] + w_0)S^T$$

Steady-state Case

The steady-state moments and covariances can be obtained by solving linear algebraic equations:

Let $\bar{X} = \lim_{t\to\infty} E[X(t)]$ and $\bar{\Sigma} = \lim_{t\to\infty} \Sigma(t)$.

Then

$$SW\bar{X} = -Sw_0$$

$$SW\bar{\Sigma} + \bar{\Sigma}W^T S^T + S\ diag(W\bar{X} + w_0)S^T = 0$$

Fluctuations Arise from Noise Driven Dynamics

Define $A = SW$, and $B = S\sqrt{diag(W\bar{X} + w_0)}$.
The steady-state covariances equation

$$SW\bar{\Sigma} + \bar{\Sigma}W^T S^T + S\ diag(W\bar{X} + w_0)S^T = 0$$

becomes

$$A\bar{\Sigma} + \bar{\Sigma}A^T + BB^T = 0 \qquad \text{Lyapunov Equation}$$

Moment Computations

- Affine Propensity
- Linear Noise Approximation

Linear Noise Approximation (LNA)

$$\text{Let } X^{\Omega}(t) := \frac{X(t)}{\Omega}$$

Linear Noise Approximation: $X^{\Omega}(t) \approx \Phi(t) + \frac{1}{\sqrt{\Omega}} V(t)$

$$\frac{d\Phi}{dt} = Sf(\Phi)$$

where $dV(t) = A(t)V(t)dt + B(t)dW_t$

$$A(t) = \frac{d[Sf(\Phi)]}{d\Phi}(\Phi_0(t)), \qquad B(t) := S\sqrt{diag[f(\Phi_0(t))]}$$

Multiplying $X^{\Omega}(t) \approx \bar{\Phi} + \frac{1}{\sqrt{\Omega}} V(t)$ by Ω, we get

population

$$X(t) \approx \Omega\bar{\Phi} + \sqrt{\Omega} V(t)$$

deterministic concentration **zero mean stochastic**

$$E[X(t)] = \Omega\bar{\Phi}$$

Let $\bar{\Sigma}$ be the covariance matrix of $\sqrt{\Omega} \cdot V(t)$. Then

$$\frac{d}{dt}\bar{\Sigma}(t) = A(t)\bar{\Sigma}(t) + \bar{\Sigma}(t)A^T(t) + \Omega B(t)B(t)^T$$

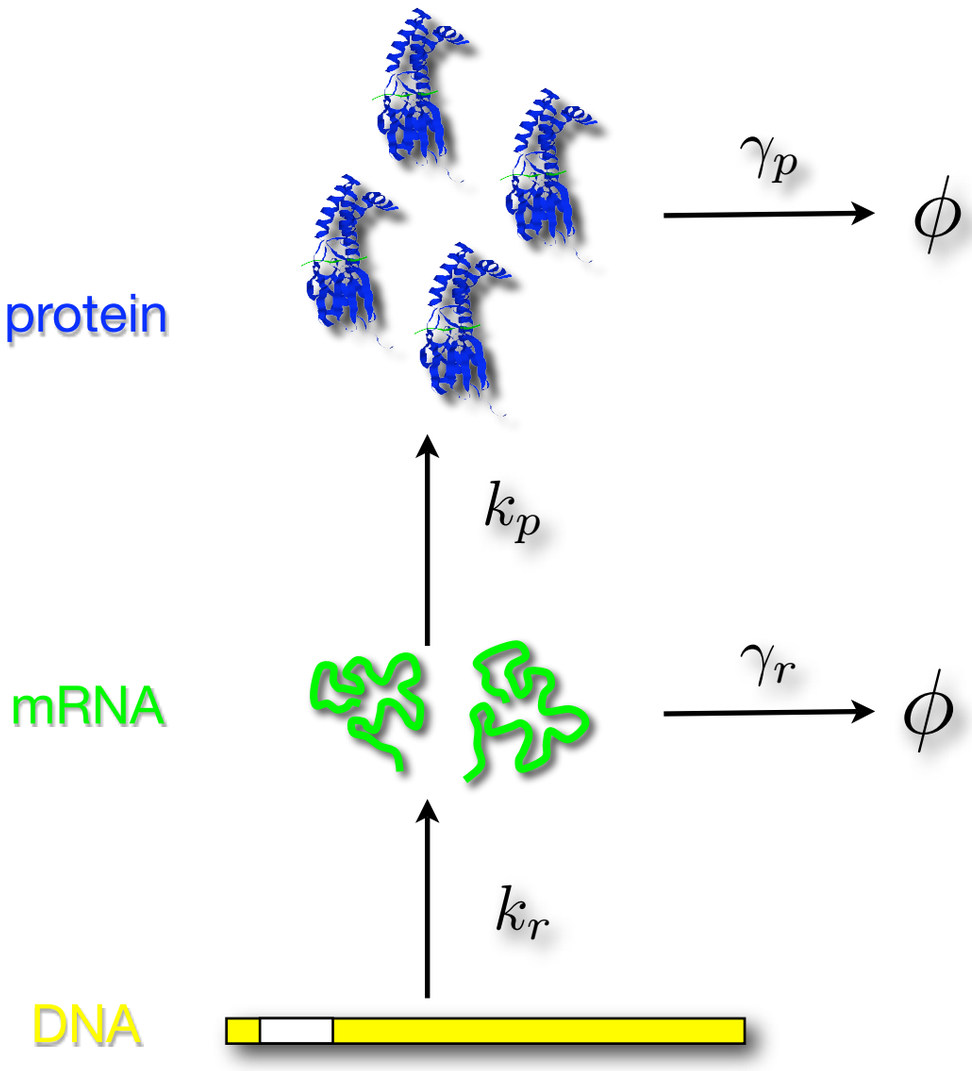
$$A(t) = \frac{d[Sf(\Phi)]}{d\Phi}(\Phi_0(t)), \qquad B(t) := S\sqrt{diag[f(\Phi_0(t))]}$$

At stationary distribution, we have the same Lyapunov equation as in the affine linear case:

$$A = SW \qquad B = S\sqrt{diag(W\bar{X} + w_0)}$$

Application to Gene Expression

Application to Gene Expression



Reactants

$X_1(t)$ is # of mRNA; $X_2(t)$ is # of protein

Reactions

$R_1 : \xrightarrow{k_r} mRNA$

$R_2 : mRNA \xrightarrow{\gamma_r}$

$R_3 : mRNA \xrightarrow{k_p} protein + mRNA$

$R_4 : protein \xrightarrow{\gamma_p} \phi$

Stoichiometry and Propensity

$$S = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

$$w(X) = \begin{bmatrix} k_r \\ \gamma_r X_1 \\ k_p X_1 \\ \gamma_p X_2 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 0 \\ \gamma_r & 0 \\ k_p & 0 \\ 0 & \gamma_p \end{bmatrix}}_{W} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} + \underbrace{\begin{bmatrix} k_r \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{w_0}$$

Steady-State Moments

$$A = SW = \begin{bmatrix} -\gamma_r & 0 \\ k_p & -\gamma_p \end{bmatrix}, \qquad Sw_0 = \begin{bmatrix} k_r \\ 0 \end{bmatrix}$$

$$\bar{X} = -A^{-1} Sw_0 = \begin{bmatrix} \frac{k_r}{\gamma_r} \\ \\ \frac{k_p k_r}{\gamma_p \gamma_r} \end{bmatrix}$$

Steady-State Covariance

$$BB^T = S\ diag(W\bar{X} + w_0)S^T = \begin{bmatrix} 2k_r & 0 \\ 0 & \frac{2k_p k_r}{\gamma_r} \end{bmatrix}$$

The steady-state covariances equation

$$A\bar{\Sigma} + \bar{\Sigma}A^T + BB^T = 0 \qquad \text{Lyapunov Equation}$$

can be solved algebraically for $\bar{\Sigma}$.

$$\bar{\Sigma} = \begin{bmatrix} \frac{k_r}{\gamma_r} & \frac{k_p k_r}{\gamma_r(\gamma_r+\gamma_p)} \\ \\ \frac{k_p k_r}{\gamma_r(\gamma_r+\gamma_p)} & \frac{k_p k_r}{\gamma_p \gamma_r}(1 + \frac{k_p}{\gamma_r+\gamma_p}) \end{bmatrix}$$

Coefficients of Variation

$$C_{vr}^2 = \frac{1}{\frac{k_r}{\gamma_r}} = \frac{1}{\bar{X}_1}$$

$$C_{vp}^2 = \frac{1}{\frac{k_r k_p}{\gamma_r \gamma_p}} \left(1 + \frac{k_p}{\gamma_r + \gamma_p}\right) = \frac{1}{\bar{X}_2} \left(1 + \frac{k_p}{\gamma_r + \gamma_p}\right)$$

Question: Does a large $\bar{X}_2$ imply a small C_{vp}?

$$\begin{aligned} C_{vp}^2 &= \frac{1}{\frac{k_r k_p}{\gamma_r \gamma_p}} \left(1 + \frac{k_p}{\gamma_r + \gamma_p}\right) \\ &\geq \frac{1}{\frac{k_r k_p}{\gamma_r \gamma_p}} \left(\frac{k_p}{\gamma_r + \gamma_p}\right) = \frac{\gamma_r \gamma_p}{k_r} \cdot \frac{1}{\gamma_r + \gamma_p} \end{aligned}$$

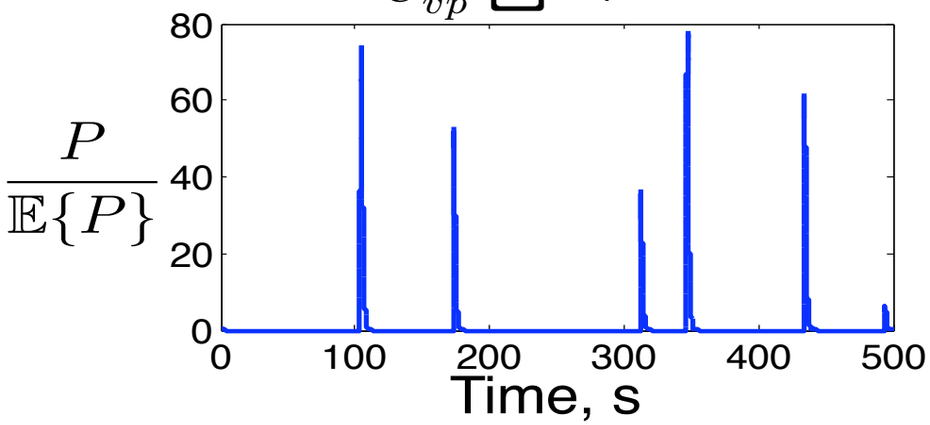
$\bar{X}_2 = \frac{k_r k_p}{\gamma_r \gamma_p}$, which can be chosen *independently* from C_{vp}.

Large mean does not imply small fluctuations!

$$\mathbb{E}\{P\} = 100, \quad \gamma_r = \gamma_p = 1$$

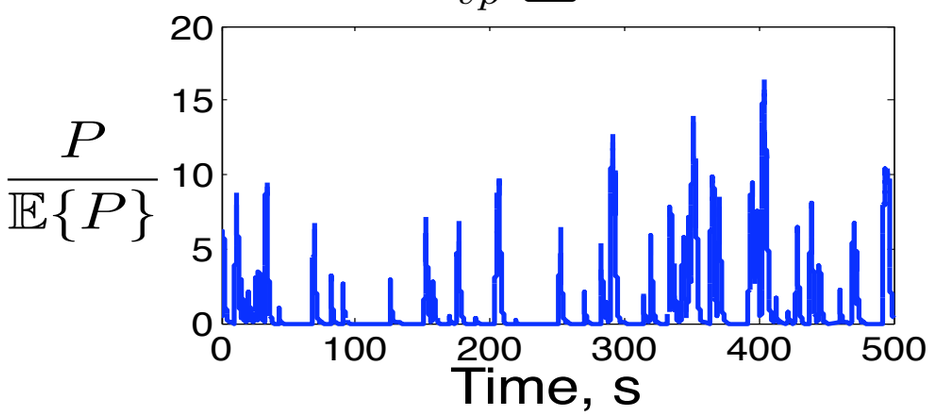
$k_r = 0.01 \quad k_p = 10,000$

C_{vp}^2 □ .



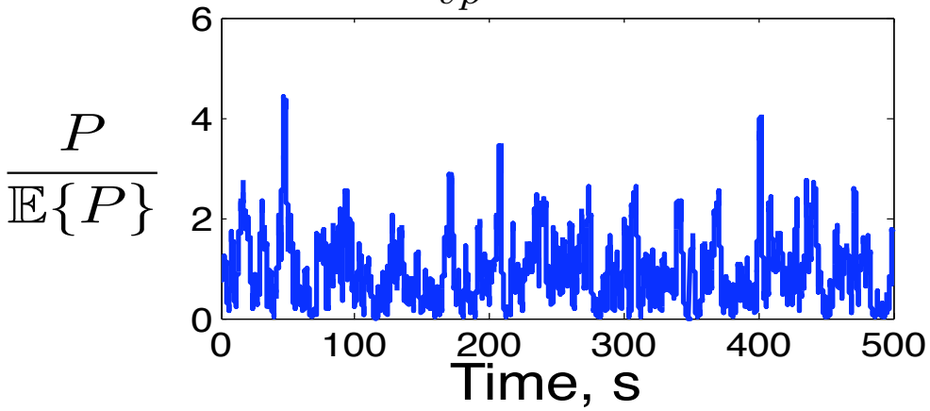
$k_r = 0.1 \quad k_p = 1000$

C_{vp}^2 □ .



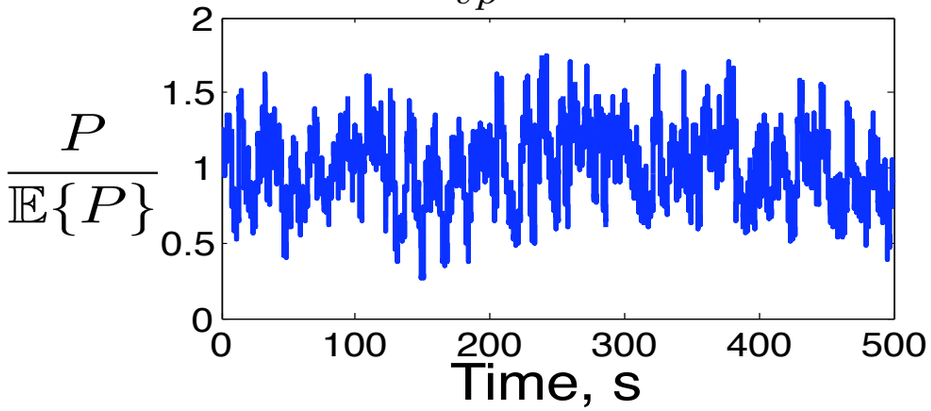
$k_r = 1 \quad k_p = 100$

$C_{vp}^2 = 0.51$



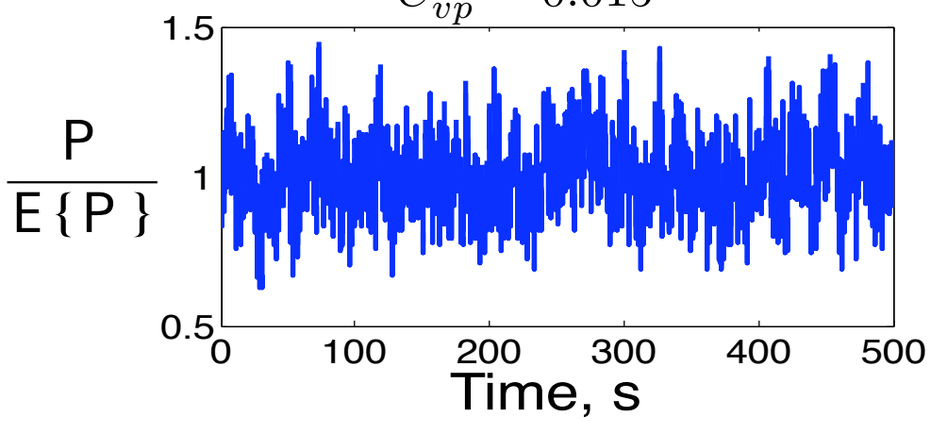
$k_r = 10 \quad k_p = 10$

$C_{vp}^2 = 0.06$



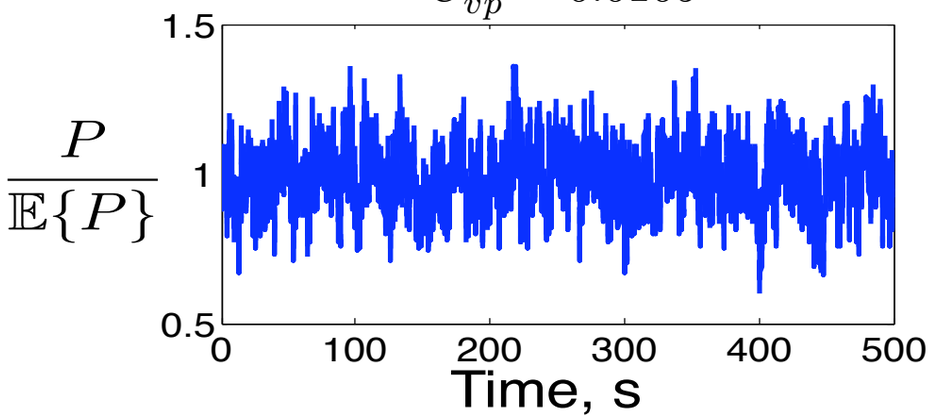
$k_r = 100 \quad k_p = 1$

$C_{vp}^2 = 0.015$



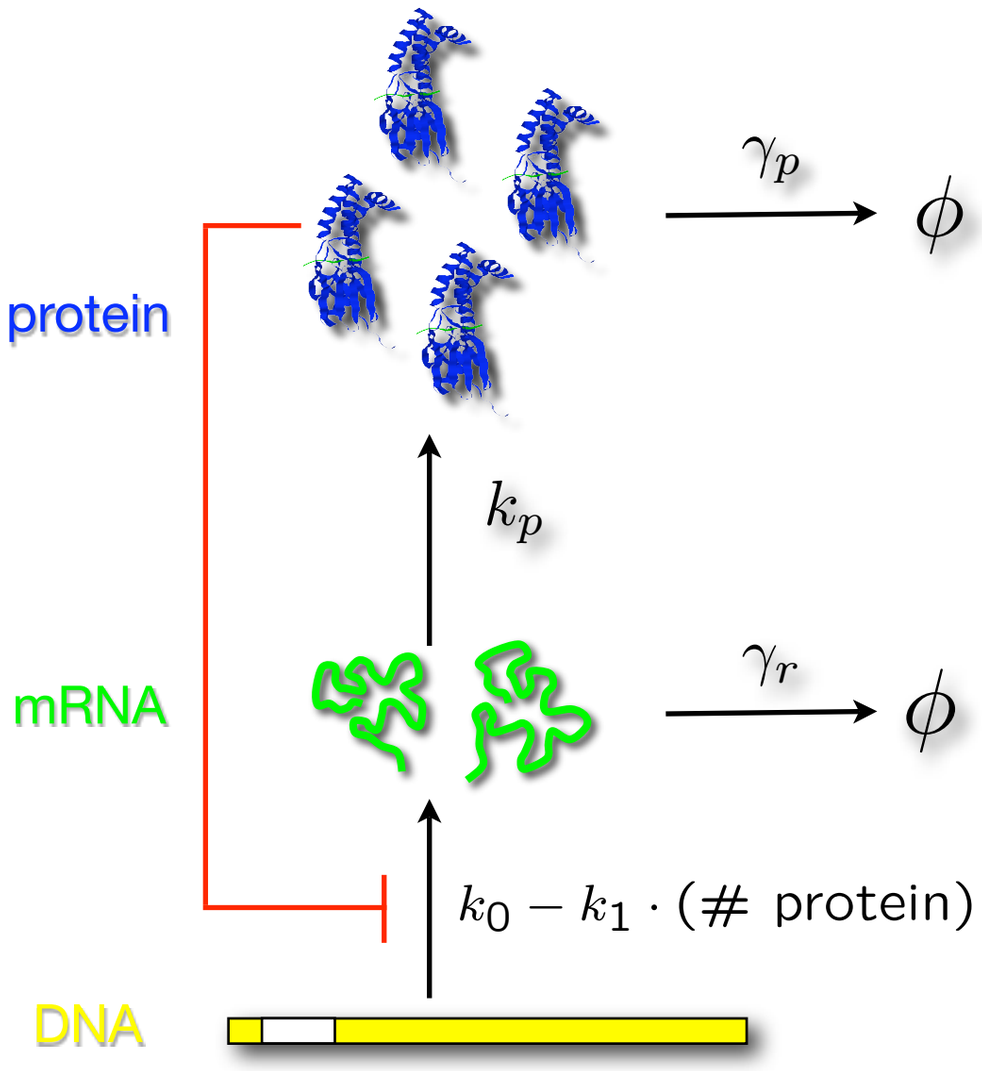
$k_r = 1000 \quad k_p = 0.1$

$C_{vp}^2 = 0.0105$



Noise Suppression and Exploitation

Noise Attenuation through Negative Feedback



Reactants

$X_1(t)$ is # of mRNA; $X_2(t)$ is # of protein

Reactions

$R_1 : \xrightarrow{k_r} mRNA$ $\qquad k_r = k_0 - k_1 \cdot (\#\ \text{protein})$

$R_2 : mRNA \xrightarrow{\gamma_r}$

$R_3 : mRNA \xrightarrow{k_p} protein + mRNA$

$R_4 : protein \xrightarrow{\gamma_p} \phi$

Stoichiometry and Propensity

$$S = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

$$w(X) = \begin{bmatrix} k_0 - k_1 X_2 \\ \gamma_r X_1 \\ k_p X_1 \\ \gamma_p X_2 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & -k_1 \\ \gamma_r & 0 \\ k_p & 0 \\ 0 & \gamma_p \end{bmatrix}}_{W} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} + \underbrace{\begin{bmatrix} k_0 \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{w_0}$$

Steady-State Moments

$$A = SW = \begin{bmatrix} -\gamma_r & -k_1 \\ k_p & -\gamma_p \end{bmatrix}, \qquad Sw_0 = \begin{bmatrix} k_0 \\ 0 \end{bmatrix}$$

$$\bar{X} = -A^{-1} Sw_0 = \begin{bmatrix} \dfrac{\frac{k_0}{\gamma_r}}{1+\frac{k_1 k_p}{\gamma_p \gamma_r}} \\ \\ \dfrac{\frac{k_0 k_p}{\gamma_r \gamma_p}}{1+\frac{k_1 k_p}{\gamma_p \gamma_r}} \end{bmatrix} =: \begin{bmatrix} \mu_r \\ \mu_p \end{bmatrix}$$

Steady-State Covariance

$$BB^T = S\ diag(W\bar{X} + w_0)S^T = \begin{bmatrix} k_0 + \ r\mu_r - k_1\mu_p & 0 \\ 0 & k_p\mu_r + \ p\mu_p \end{bmatrix}$$

The steady-state covariances equation

$$A\bar{\Sigma} + \bar{\Sigma}A^T + BB^T = 0 \qquad \text{Lyapunov Equation}$$

can be solved algebraically for $\bar{\Sigma}$.

$$\bar{\Sigma}_{22} = \sigma_p^2 = \left[\frac{1-\phi}{1+b\phi} \cdot \frac{b}{1+\eta} + 1\right] \mu_p \qquad \text{where } \phi = \frac{k_1}{\gamma_p},\ b = \frac{k_p}{\gamma_r},\ \eta = \frac{\gamma_p}{\gamma_r}$$

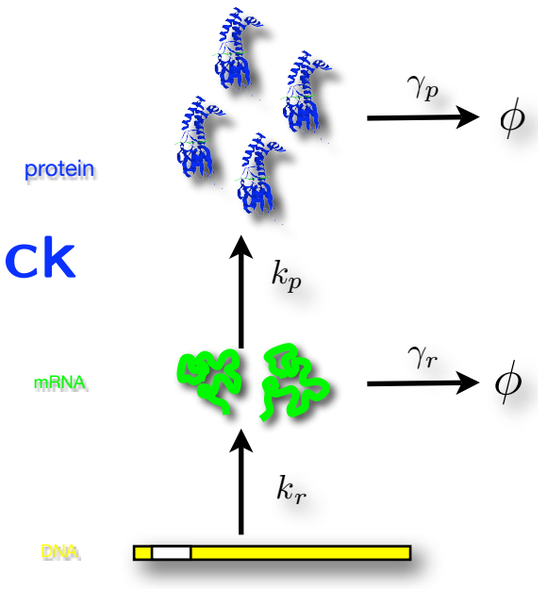
Feedback vs. No Feedback

In order to compare the noise in the two cases, we must ensure that both configuations have the same mean!

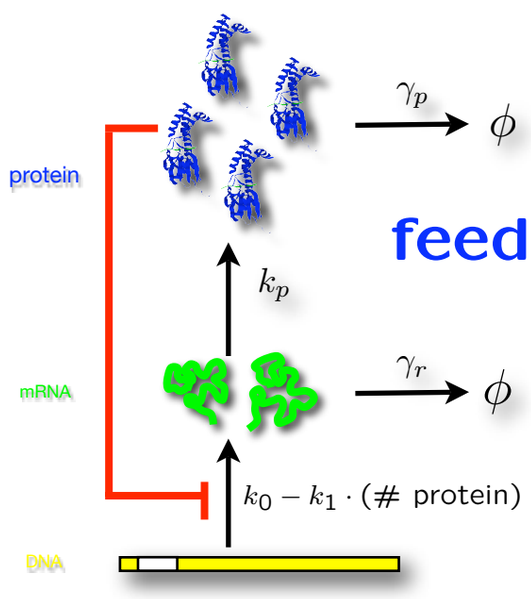
Impose the constraint: $\mu_p^{FB} = \mu_p^{NFB} =: \mu_p^*$

This may be achieved by choosing $k_0 = k_r + k_1 \mu_p^{NFB}$.

no feedback



feedback

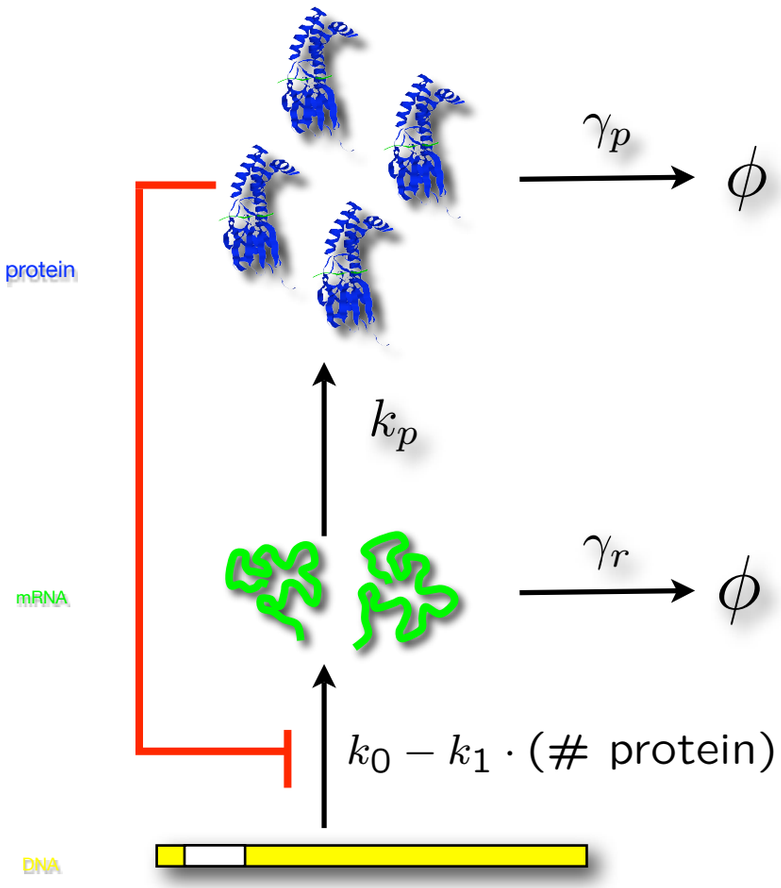


Mean	μ_p^*	μ_p^*
Variance	$\left[\frac{b}{1+\eta} + 1\right] \mu_p^*$	$\left[\frac{1-\phi}{1+b\phi} \cdot \frac{b}{1+\eta} + 1\right] \mu_p^*$ < 1

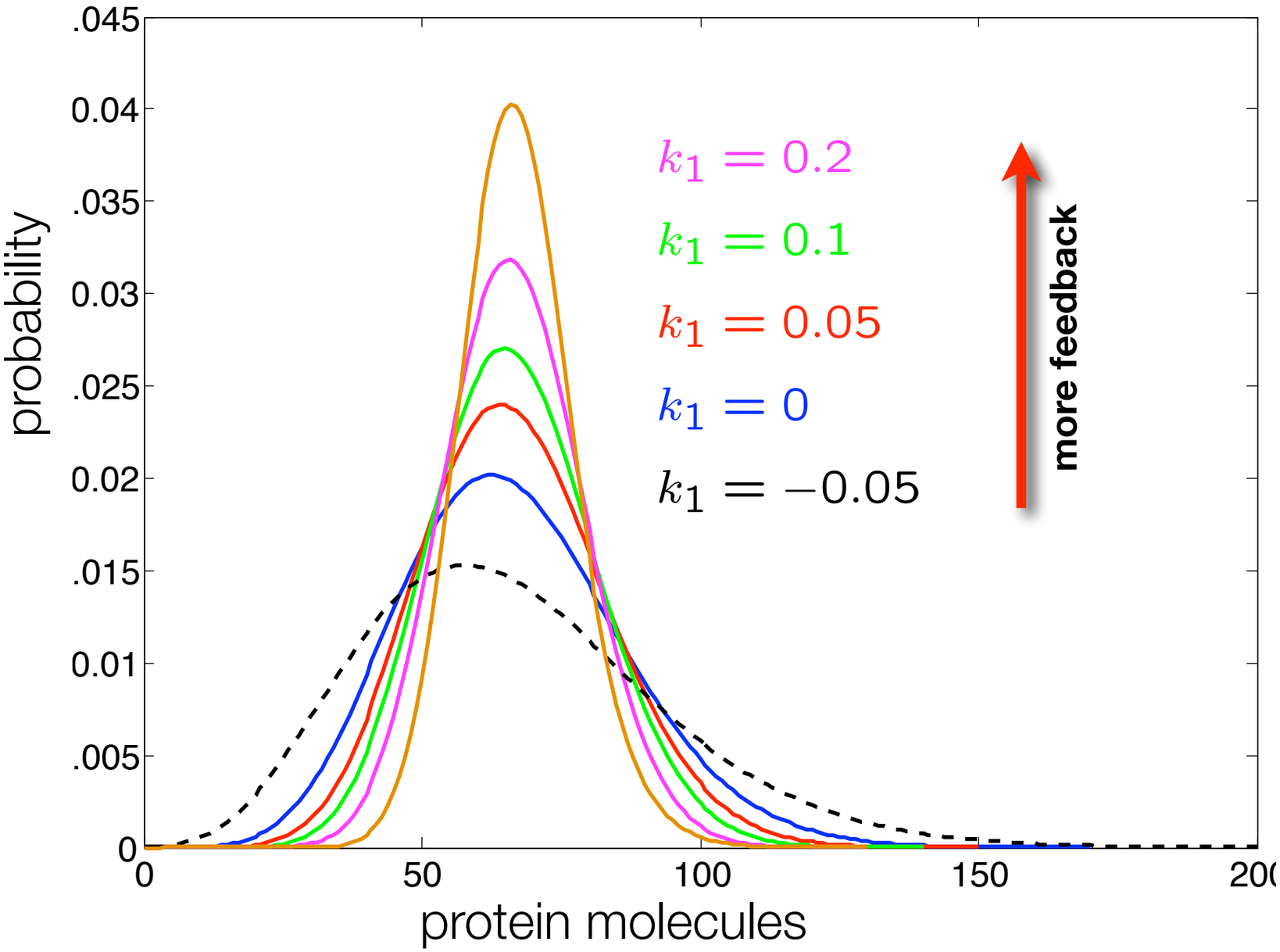
where $\phi = \frac{k_1}{\gamma_p}$

Protein variance is always smaller with negative feedback!

Example



$\gamma_p = \gamma_r = 1 \qquad k_p = 10;$



Note that these distributions are NOT Gaussian.

Exploiting the Noise: Failure of the linear noise approximation

$$\phi \underset{k_a S}{\overset{k}{\rightleftharpoons}} I \xrightarrow{k_p} P \xrightarrow{1} \phi$$

$$\phi \underset{k_d}{\overset{k_s}{\rightleftharpoons}} S$$

may be approximated by

$$\phi \xrightarrow{kq} P \xrightarrow{1} \phi$$

$$q = \frac{1}{1 + \frac{n}{\Omega K}} \qquad K = k_p/k_a$$

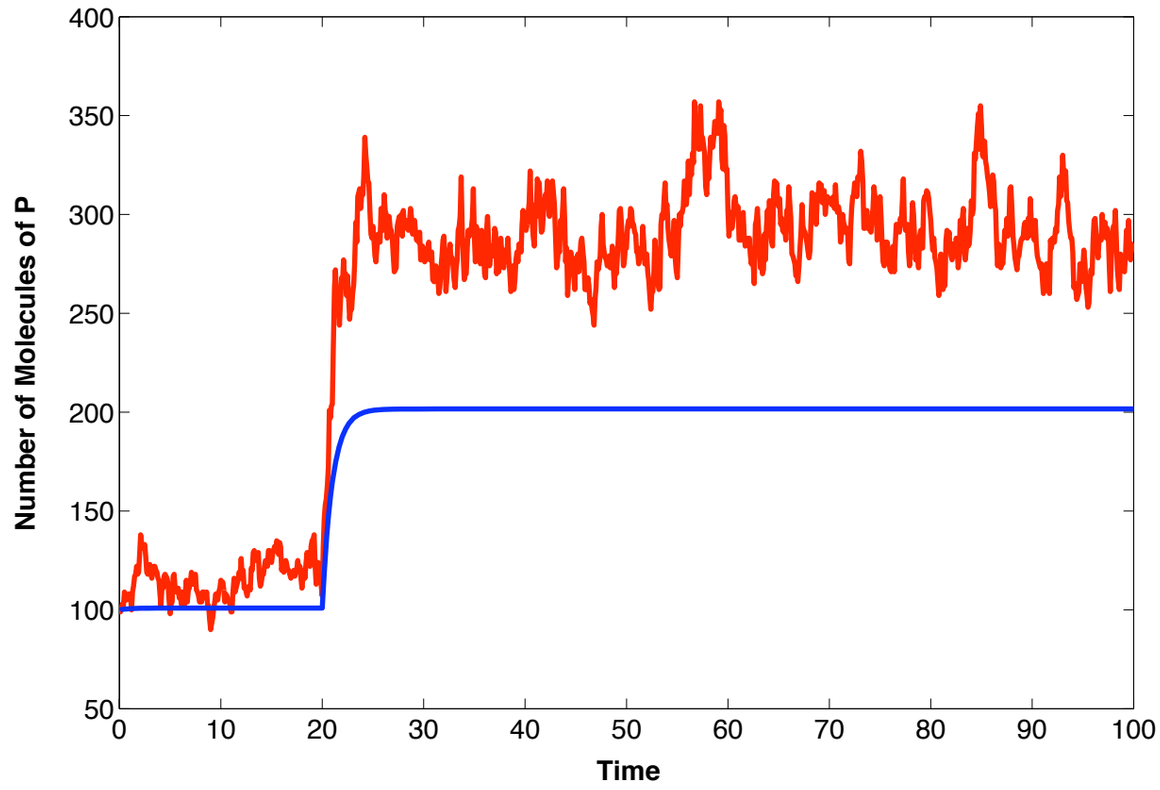
n is #S

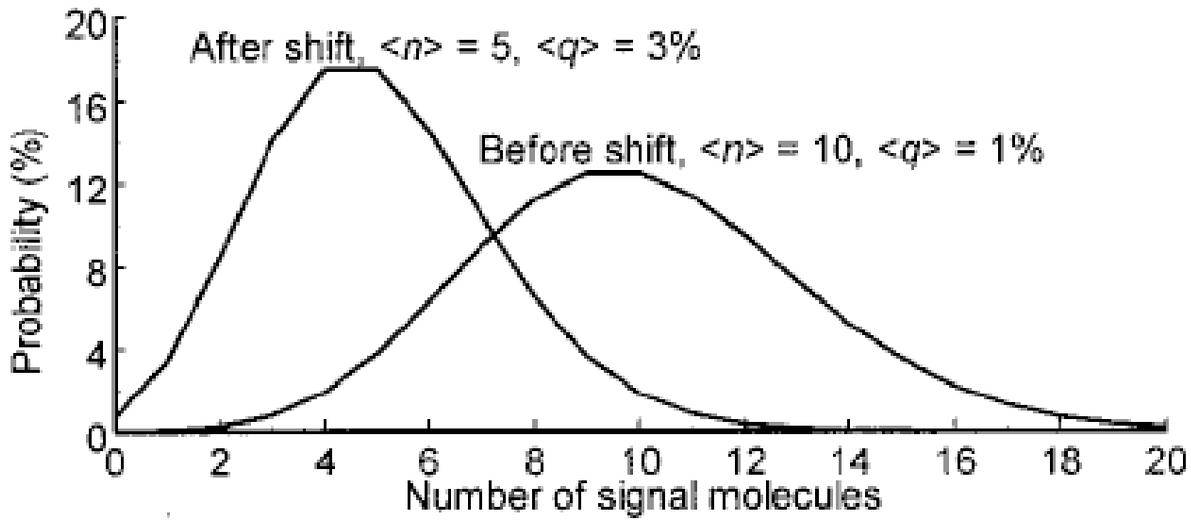
convex

From Jensen's Inequality:

$$E[q] = E\left[\frac{1}{1 + \frac{n}{\Omega K}}\right] \geq \frac{1}{1 + \frac{E[n]}{\Omega K}}$$

- Noise *enhances* signal!





Johan Paulsson , Otto G. Berg , and Måns Ehrenberg, PNAS 2000

On the menu...

- Today
  - ☑ Overview of Stochastic Gene Expression
  - ☑ Stochastic Chemical Kinetics
  - ☑ Solutions for Simple Stochastic Processes (Transcription)
  - ☑ Importance of Population Size
  - ☑ Moment Computations for Linear Propensities
  - ☑ Linear Noise Approximation
- Tomorrow
  - ▸ Monte Carlo Simulation Techniques
    - ✳ Gillespie (SSA), Tau leaping, Chemical Langevin (SDEs), Slow Scale SSA.
  - ▸ Density Computations with Finite State Projection Techniques
  - ▸ Switch and Trajectory Analyses

Monte Carlo Methods

Kinetic Monte-Carlo Simulation Methods

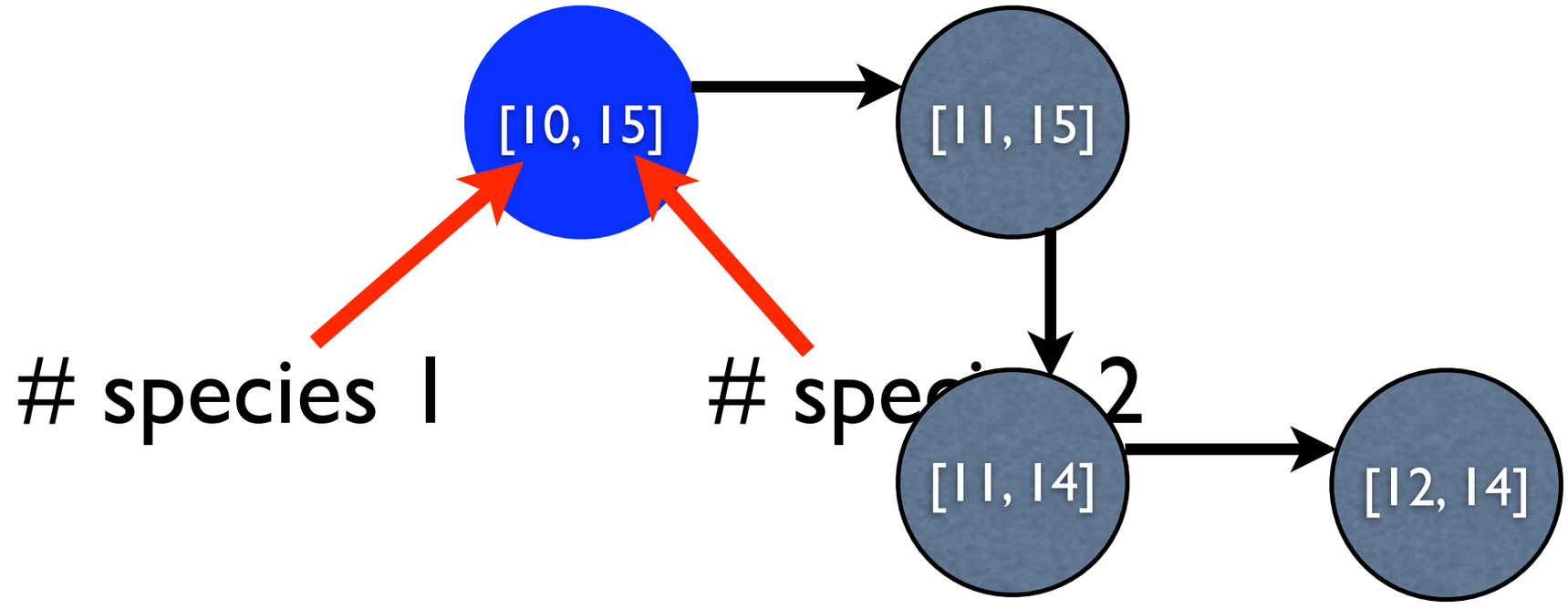
Brian Munsky

- Stochastic Simulation Algorithm
  - D.T. Gillespie, J. Phys. Chem. A **81**, 2340 (1977)
  - M. Gibson and J. Bruck, J. Phys. Chem. **104**, 1876 (2000)
- τ leaping
  - D. Gillespie, J. Chem. Phys. **115**, 1716 (2001); **119**, 8229 (2003)
  - M. Rathinam *et al.*, J. Chem. Phys. **119**, 12784 (2003)
  - T. Tian and K. Burrage, J. Chem. Phys. **121**, 10356 (2004)
  - A. Chatterjee, et al. J. Chem. Phys. **122**, 054104 (2005)
  - Y. Cao, D. Gillespie and L. Petzold, J. Chem. Phys. **123**, 054104 (2005)
- Chemical Langevin Equations
  - D. Gillespie, J. Chem. Phys. **113**, 1716 (2000)
- System Partitioning Methods
  - C. Rao and A. Arkin, J. Chem. Phys. **118**, 4999 (2003)
  - Y. Cao *et al.*, J. Chem. Phys. **122**, 014116 (2005)
- Hybrid Methods
  - E. Haseltine and J. Rawlings, J. Chem. Phys. **117**, 6959 (2002)
  - H. Salis and Y. Kaznessis, J. Chem. Phys. **122**, 054103 (2005)

66

A Jump-Markov description of chemical kinetics

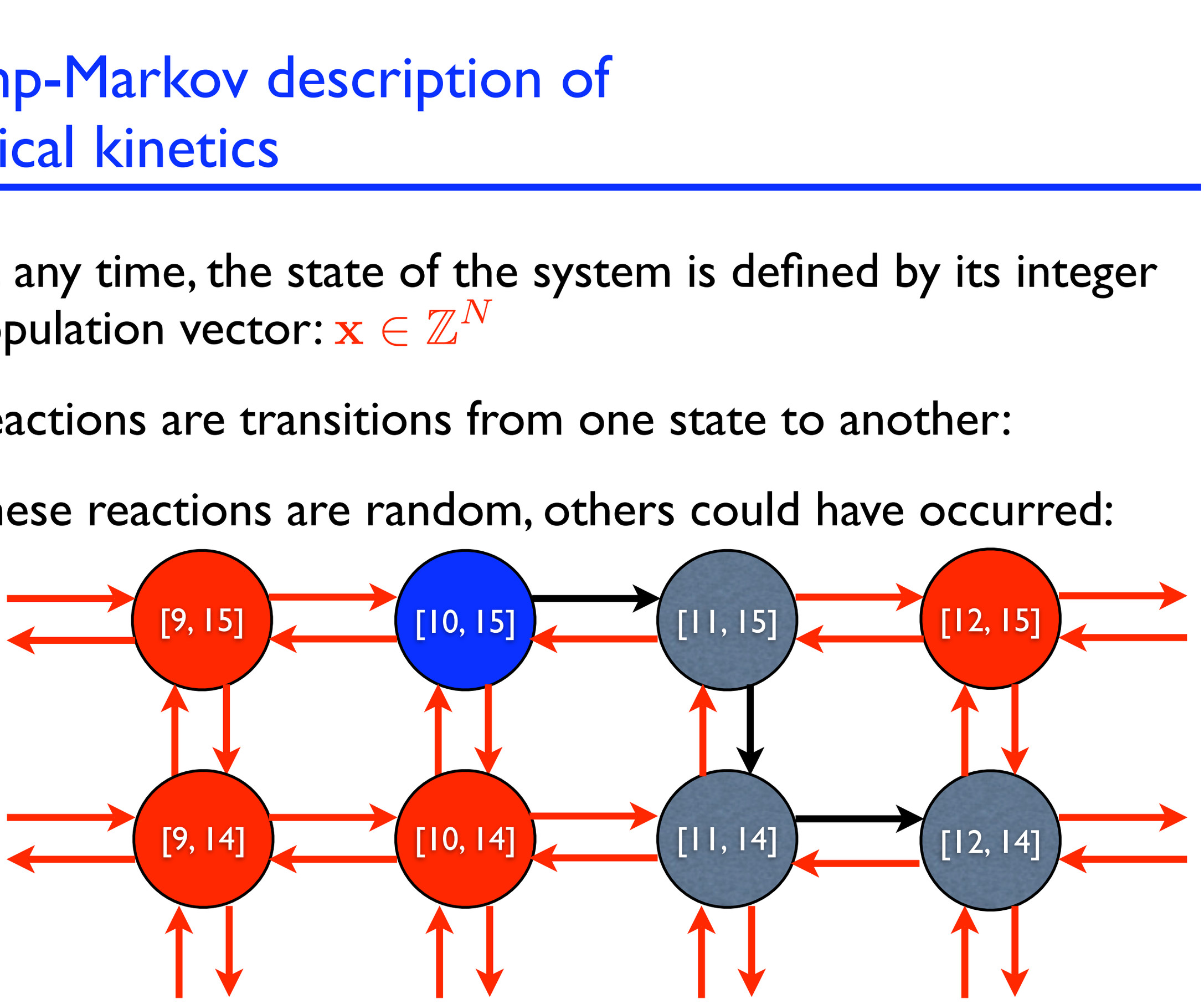
- At any time, the state of the system is defined by its integer population vector: $\mathbf{x} \in \mathbb{Z}^N$
- Reactions are transitions from one state to another:



67

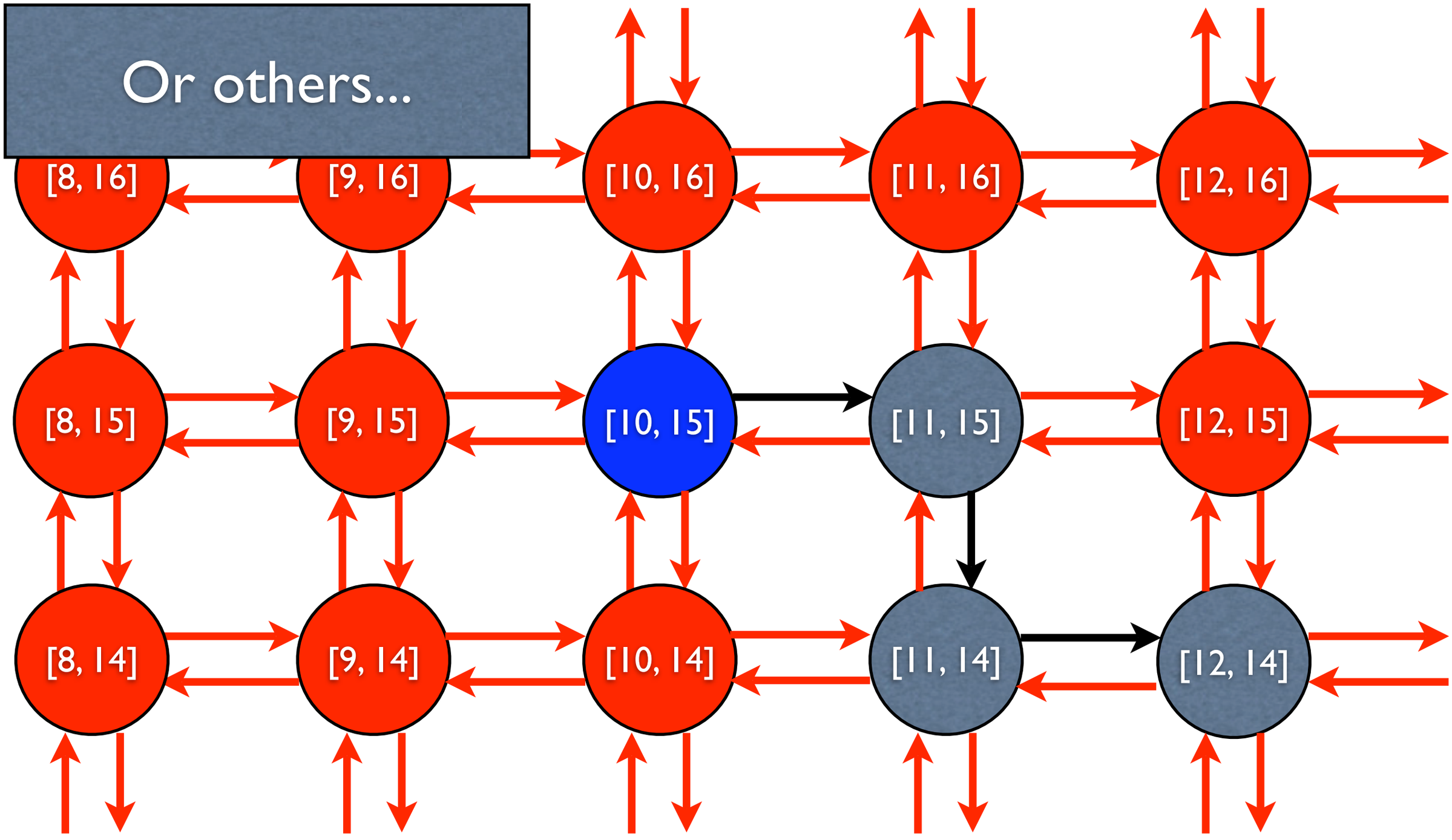
A Jump-Markov description of chemical kinetics

- At any time, the state of the system is defined by its integer population vector: $\mathbf{x} \in \mathbb{Z}^N$
- Reactions are transitions from one state to another:
- These reactions are random, others could have occurred:



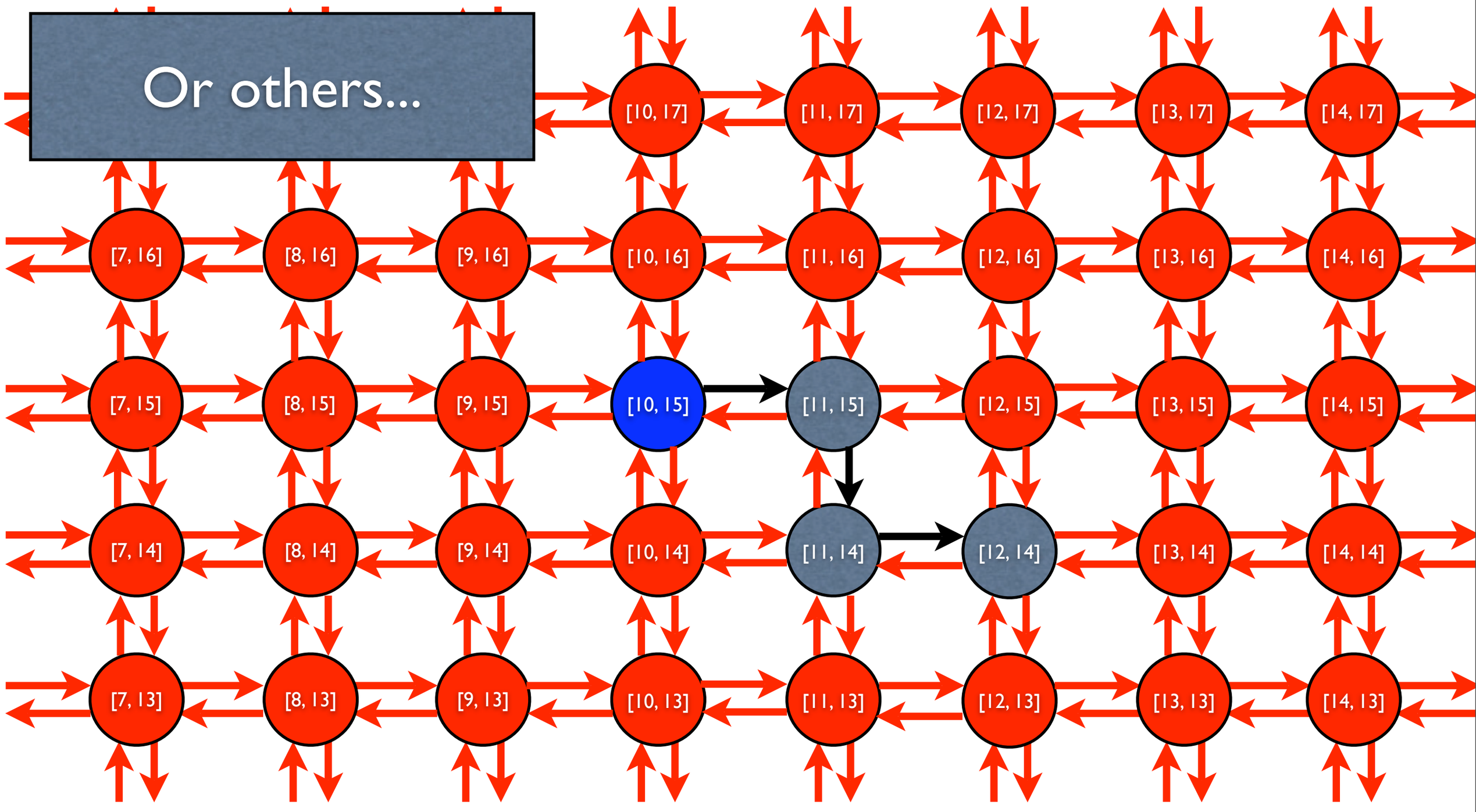
68

A Jump-Markov description of chemical kinetics



69

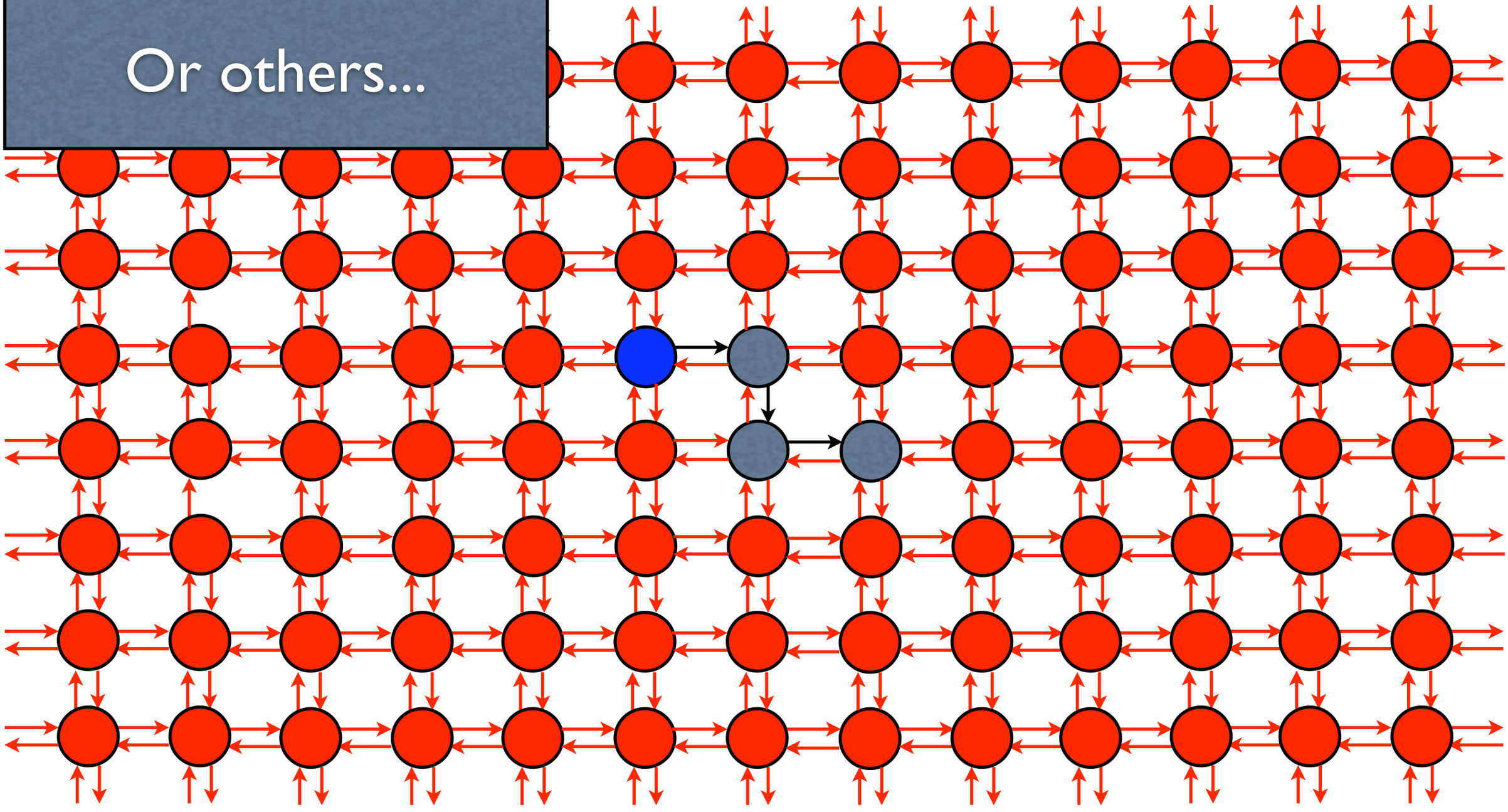
A Jump-Markov description of chemical kinetics



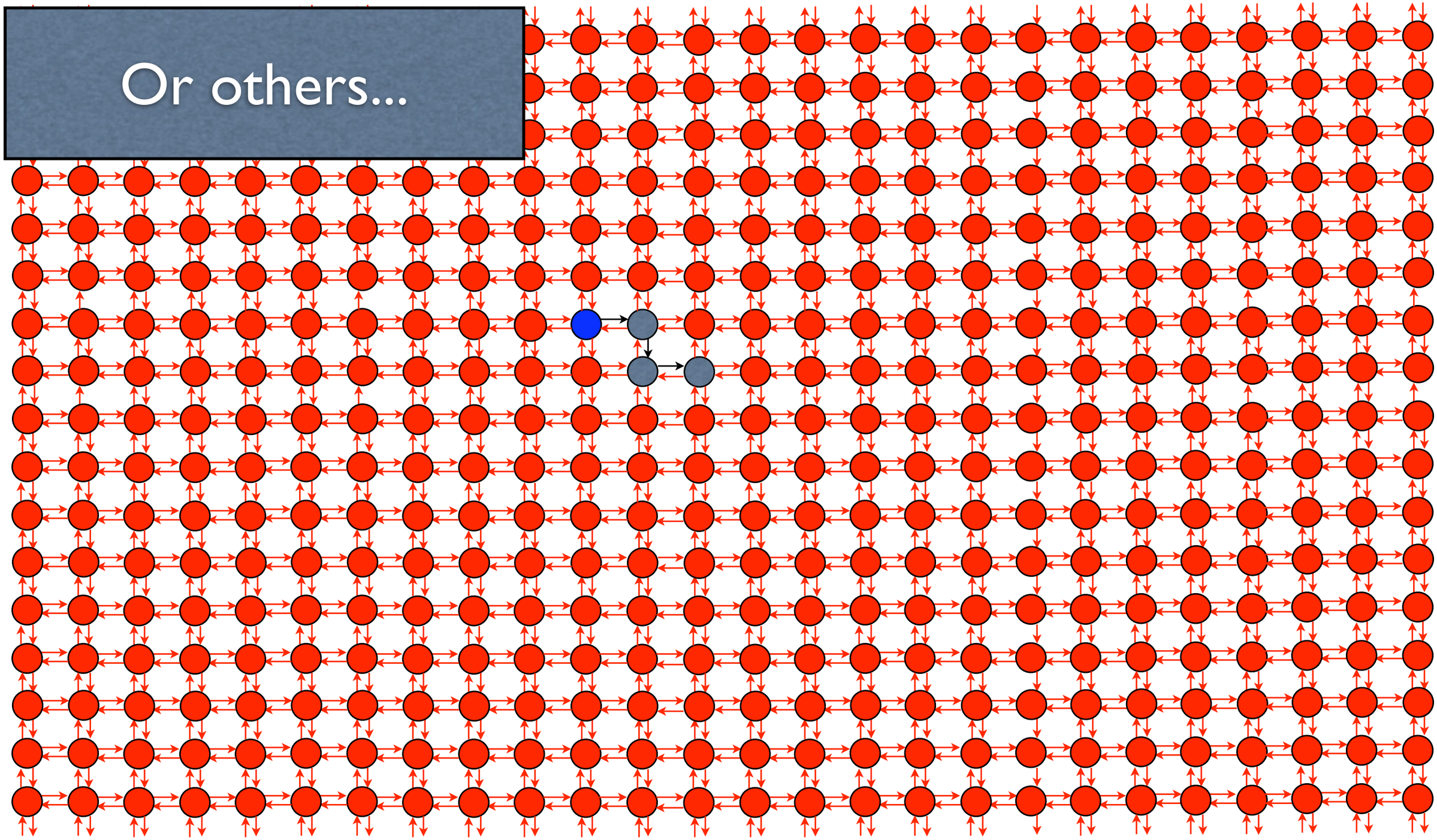
70

A Jump-Markov description of chemical kinetics

Or others...



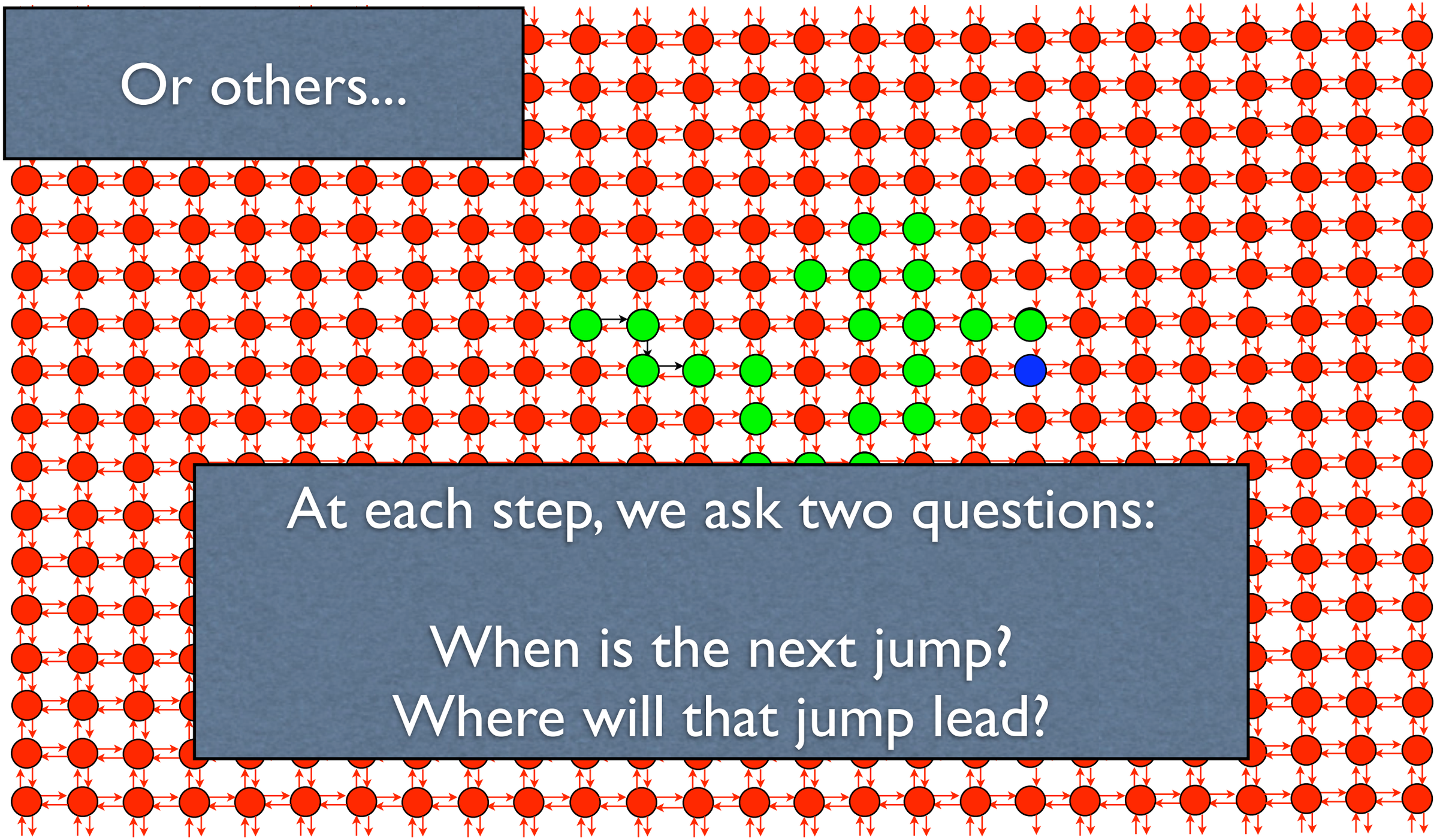
A Jump-Markov description of chemical kinetics



72

A Jump-Markov description of chemical kinetics

Or others...



At each step, we ask two questions:

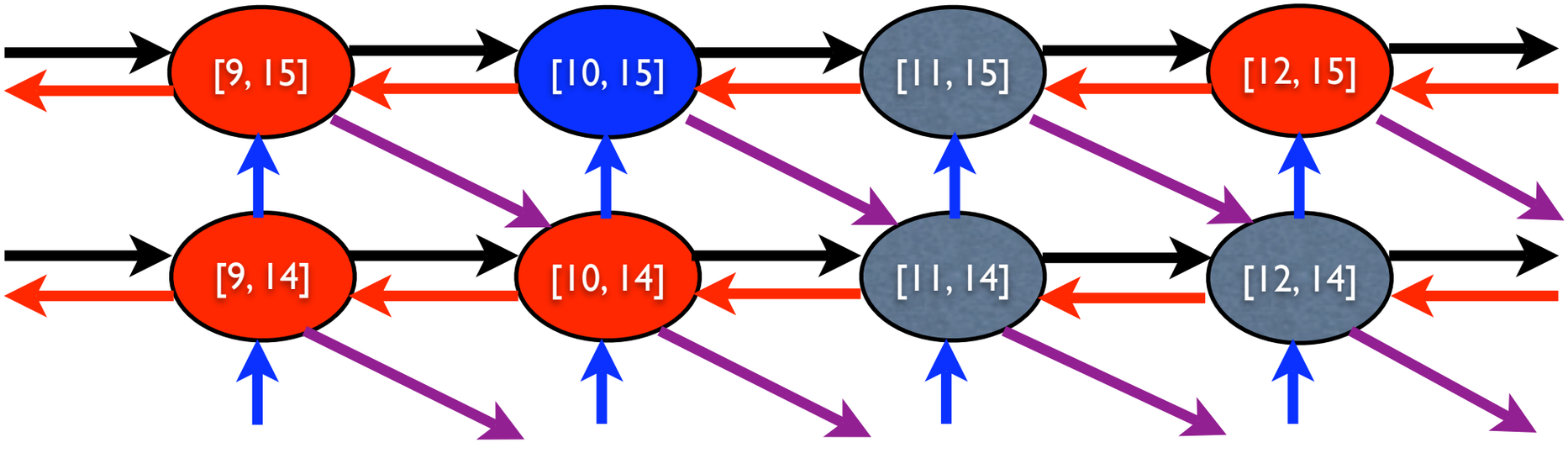
When is the next jump?
Where will that jump lead?

73

Reaction Stoichiometry (review)

- The Stoichiometric vector, **s**, refers to the relative change in the population vector after a reaction.
- There may be many different reactions for a given stoichiometry.

$\mathbf{s}_1 = [1, 0]^T$	$\mathbf{s}_2 = [-1, 0]^T$	$\mathbf{s}_3 = [0, 1]^T$	$\mathbf{s}_4 = [1, -1]^T$
$\mathcal{S}_1 \rightarrow \mathcal{S}_1 + \mathcal{S}_1$	$\mathcal{S}_1 + \mathcal{S}_1 \rightarrow \mathcal{S}_1$	$\mathcal{S}_2 \rightarrow \mathcal{S}_2 + \mathcal{S}_2$	$\mathcal{S}_2 \rightarrow \mathcal{S}_1$
$\mathcal{S}_2 \rightarrow \mathcal{S}_2 + \mathcal{S}_1$	$\mathcal{S}_1 + \mathcal{S}_2 \rightarrow \mathcal{S}_2$	$\mathcal{S}_1 \rightarrow \mathcal{S}_1 + \mathcal{S}_2$	$\mathcal{S}_1 + \mathcal{S}_2 \rightarrow \mathcal{S}_1 + \mathcal{S}_1$
$\emptyset \rightarrow \mathcal{S}_1$	$\mathcal{S}_1 \rightarrow \emptyset$	$\emptyset \rightarrow \mathcal{S}_2$	$\mathcal{S}_2 + \mathcal{S}_2 \rightarrow \mathcal{S}_1 + \mathcal{S}_2$

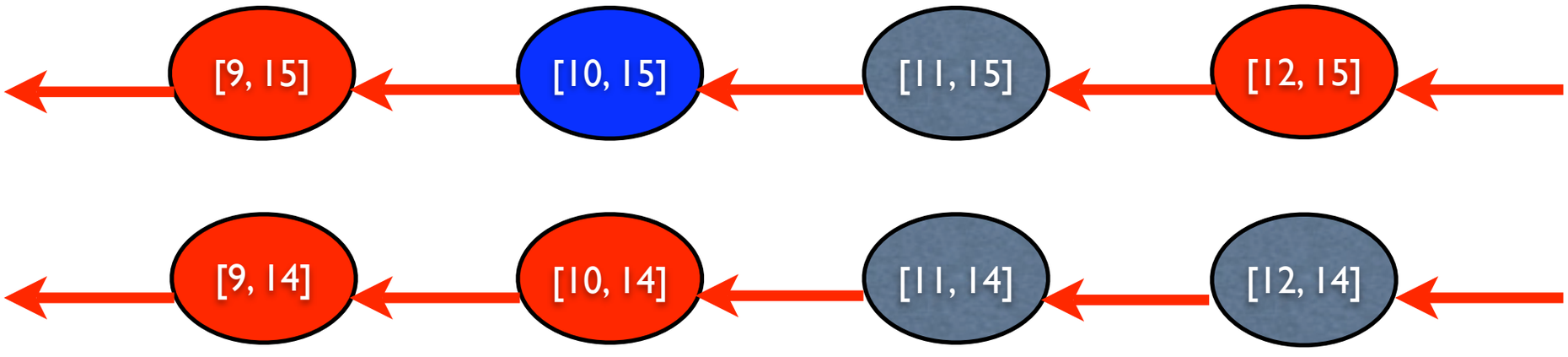


74

Reaction Propensities (review)

- The propensity, w, of a reaction is its rate.
- $\mathrm{w}_\mu dt$ is the probability that the μ^{th} reaction will occur in a time step of length dt.
- Typically, propensities depend only upon reactant populations.

$\mathbf{s}_2 = [-1, 0]^T$	$w_2(x_1, x_2)$
$\mathcal{S}_1 + \mathcal{S}_1 \rightarrow \mathcal{S}_1$	$k_1 x_2 (x_1 - 1)/2$
$\mathcal{S}_1 + \mathcal{S}_2 \rightarrow \mathcal{S}_2$	$k_2 x_1 x_2$
$\mathcal{S}_1 \rightarrow \emptyset$	$k_3 x_1$



75

Exponential Waiting Times

Probability reaction will occur in $[t, t+\Delta t)$: $w\Delta t + \mathcal{O}(\Delta t)^2$

Probability reaction will *not* occur in $[t, t+\Delta t)$: $1 - w\Delta t + \mathcal{O}(\Delta t)^2$

Probability a reaction will not occur in two such time intervals $[t, t+2\Delta t)$: $\left(1 - w\Delta t + \mathcal{O}(\Delta t)^2\right)^2 = 1 - 2w\Delta t + \mathcal{O}(\Delta t)^2$

Suppose that, $\tau = K\Delta t$, then the probability that no reaction will occur in the interval $[t, t+\tau)$ is

$$\left(1 - w\frac{\tau}{K} + \mathcal{O}(K^{-2})\right)^K$$

Taking the limit as K goes to infinity yields that the probability that no reaction will occur in the interval $[t, t+\tau)$ is

$$\lim_{k\to\infty} \left(1 - w\frac{\tau}{K} + \mathcal{O}(K^{-2})\right)^K = \exp(-w\tau)$$

Exponential Random Variables

The probability that a reaction will occur in the interval $[t, t+\tau)$ is $F_T(\tau) = 1 - \exp(-w\tau)$. This is a cumulative distribution.

The density (derivative) of the random number, T, is:

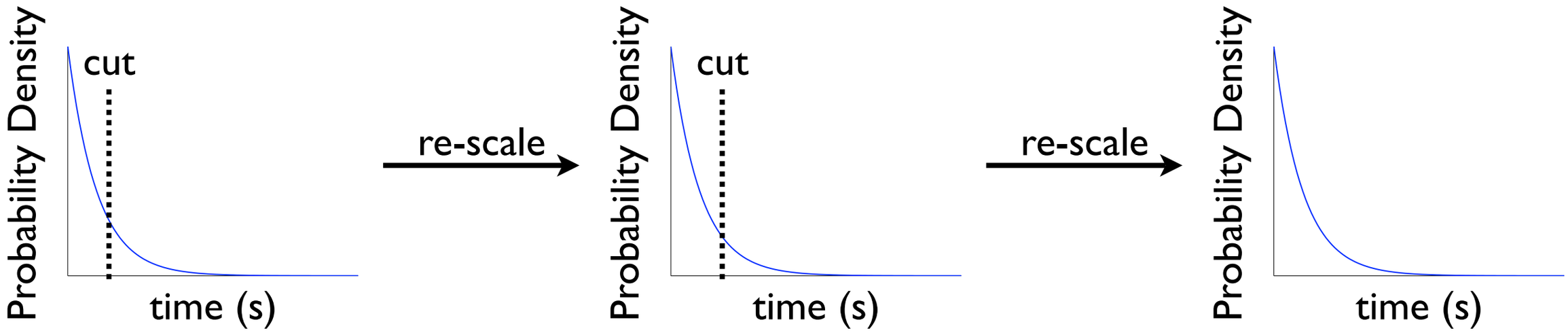
$$f_T(\tau) = \frac{1}{w} \exp(-w\tau)$$

Such a random number is known as an *exponentially* distributed random number.

Notation: $T \in \mathrm{EXP}(\lambda) \rightarrow$ T is an exponentially distributed r.v. with parameter: λ .

Exponential Waiting Times

- We have assumed that the system is fully described by the population vectors.
- If no reaction occurs, then nothing will have changed.
- Waiting times must be *memoryless* random variables.



- No matter where we cut and scale the distribution, it must always looks the same.

The exponential is the *only* continuous r.v. with this property.

78

Generating Waiting Times

- To generate an exponentially distributed random number, all we need is a uniform random number generator.

- Find the cumulative distribution,

$$F(t) = 1 - \exp(-\lambda t)$$

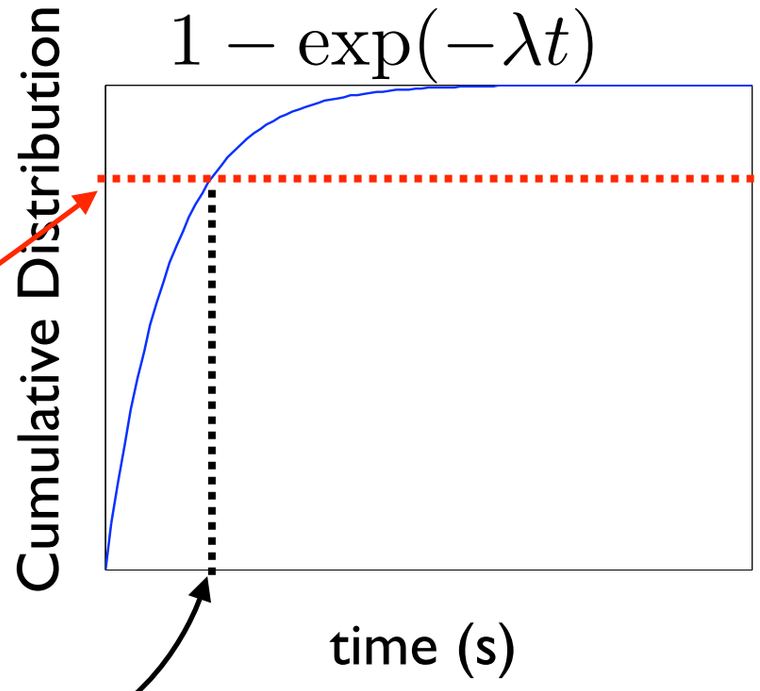
- Generate uniform random number,

$$r \in \mathrm{U}[0, 1]$$

- Find intersection where $F(t) = r$:

$$\tau = \frac{1}{\lambda} \log \frac{1}{1 - r}$$

- This is the time of the next reaction.



79

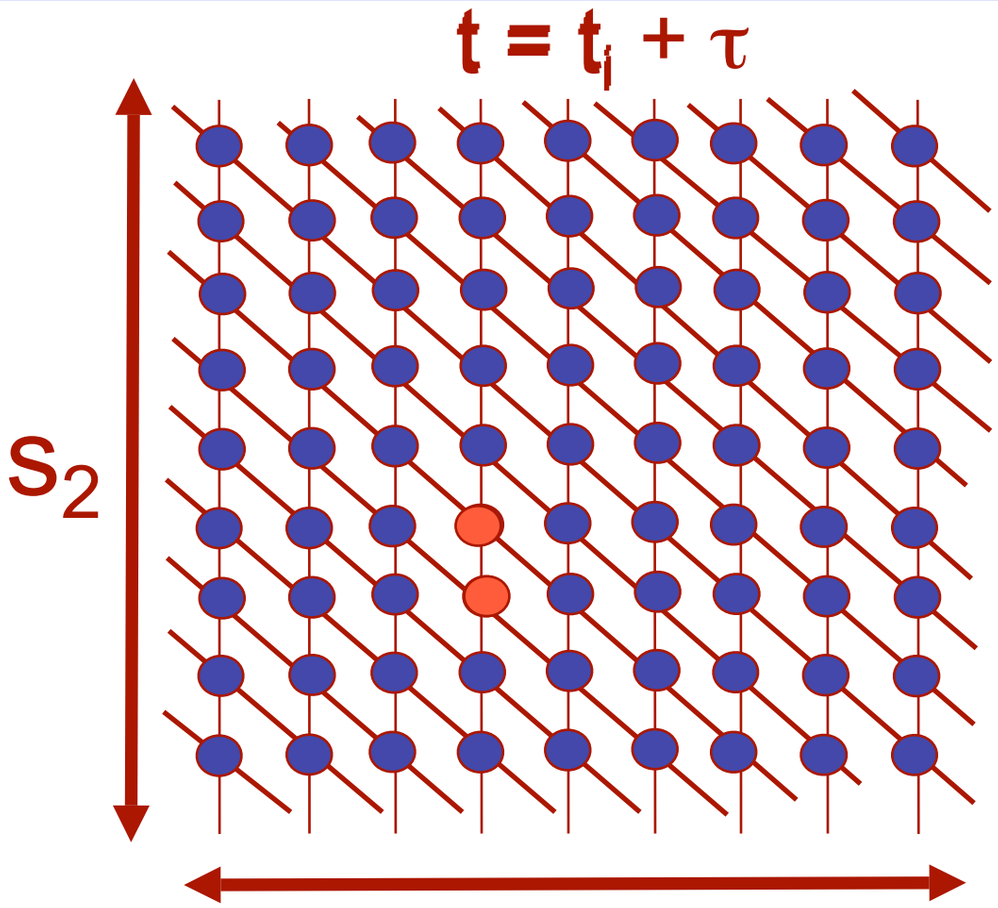
Monte-Carlo Simulation Methods

☑ The Jump Markov Process

- **Stochastic Simulation Algorithm**
  - D.T. Gillespie, J. Phys. Chem. A **81**, 2340 (1977)
  - M. Gibson and J. Bruck, J. Phys. Chem. **104**, 1876 (2000)

80

Stochastic Simulation Algorithm



Step 1. Generate the time of the next reaction.

Step 2. Decide which reaction has occurred.

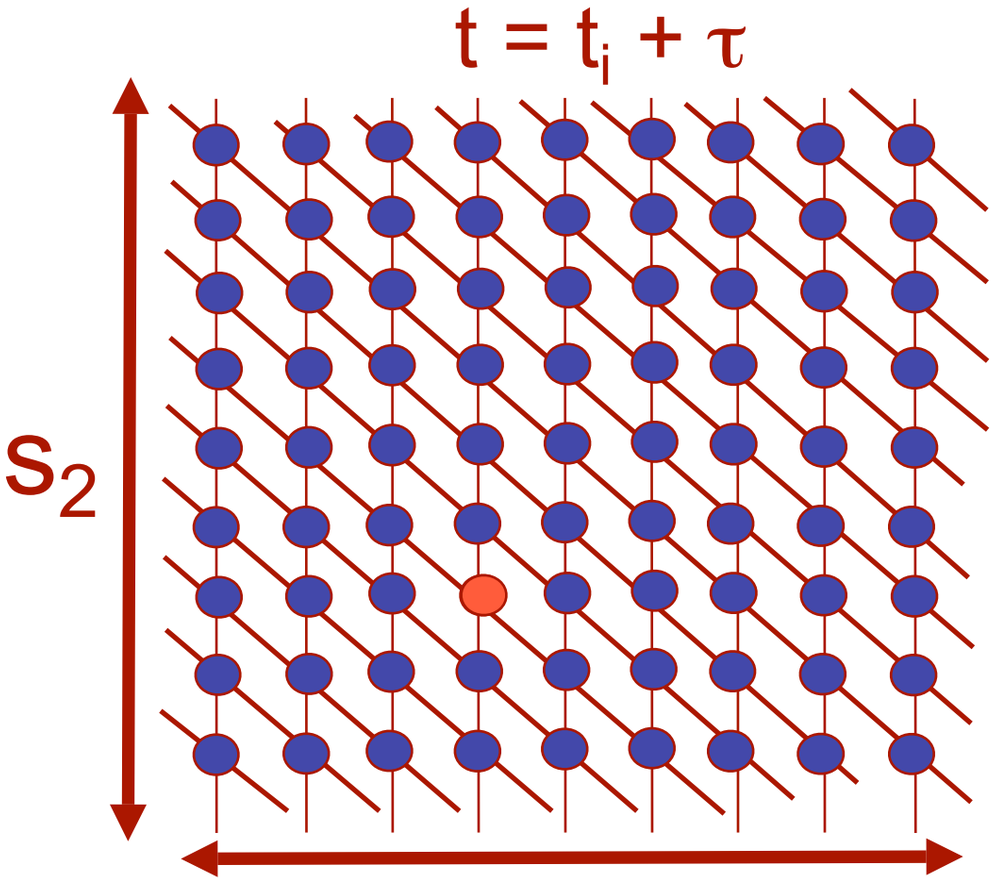
Step 3. Update current Time ($t=t+\tau$) and State ($\mathbf{x} = \mathbf{x}+s_k$).

Monte-Carlo Simulation Methods

- Stochastic Simulation Algorithm
  - D.T. Gillespie, J. Phys. Chem. A **81**, 2340 (1977)
  - M. Gibson and J. Bruck, J. Phys. Chem. **104**, 1876 (2000)
- Possible SSA methods:
  - First Reaction Method (Gillespie ‘77)
    - Next Reaction Method (Gibson and Bruck ‘00)
  - Direct Method (Gillespie ‘77)

82

The First Reaction Method (FRM)



Step 1. Generate the time of the next reaction of each type.

The time until the next reaction is a random variable of exponential distribution:

$$\tau_\mu \in \mathrm{EXP}\left(w_\mu(\mathbf{x})\right)$$

To generate each next reaction time, generate r_1 from a uniform distribution on (0,1) and use the equation:

$$\tau_\mu = \frac{1}{w_\mu(\mathbf{x})} \log \frac{1}{r_1}$$

Step 2. Decide which reaction has occurred.

This is simply the reaction with the smallest τ_μ:

$$k = \arg \left\{ \min_{\mu \in \{0,\ldots,M\}} \tau_\mu \right\}$$

Step 3. Update current Time (t=t+τ_k) and State (**x** = **x**+$\mathbf{s}_k$).

In the FRM each reaction requires M rv's.

83

The First Reaction Method
SSA in Matlab.

```
clear all
t=0;tstop = 2000;                                   %%specify initial and final times
x = [0; 0];                                         %% Specify initial conditions
S = [1 -1 0  0; 0  0 1 -1];                         %% Specify stoichiometry
w = inline('[10, 1*x(1), 10*x(1), 1*x(2)]','x');    %% Specify Propensity functions
while t<tstop
    tpos = 1./w(x).*log(1./rand(4,1));             % possible times until first reaction
    [tpos,i]=min(tpos);                            % find which is first reaction
    t=t+tpos;
    if t<=t_stop
        x = x+S(:,i);                               % update the configuration
    end
end
```


84

The Next Reaction Method (NRM)

- In the FRM, we generate times, $\{\tau_\mu\}$, for all *M* reactions and choose the reaction, *k*, with the smallest time, τ_k.
- Only a few species will change population as a result of this reaction--the rest will remain constant.
- For most reactions, the propensity functions will remain constant.
  - For these, the times can be reused in the subsequent step to find the next reaction:$\{\tau_\mu\} \rightarrow \{\tau_\mu - \tau_k\}$.
- When there are many different species and reactions, this NRM approach can be done with far fewer random number than the FRM.
- Particularly useful for compartmental or Reaction-Diffusion processes.

85

Monte-Carlo Simulation Methods

- Stochastic Simulation Algorithm
  - D.T. Gillespie, J. Phys. Chem. A **81**, 2340 (1977)
  - M. Gibson and J. Bruck, J. Phys. Chem. **104**, 1876 (2000)
- Possible SSA methods:
  - First Reaction Method (Gillespie ‘77)
  - Next Reaction Method (Gibson and Bruck ‘00)
- Direct Method (Gillespie ‘77)

86

Minimum of two Exponential Random Variables

Let $\{\tau_1, \tau_2, \ldots, \tau_M\}$ be a set of exponentially distributed random variables: $\tau_\mu \in \mathrm{EXP}\,(w_\mu)$

The minimum of $\{\tau_\mu\}$ is an exponentially distributed random variable given by:

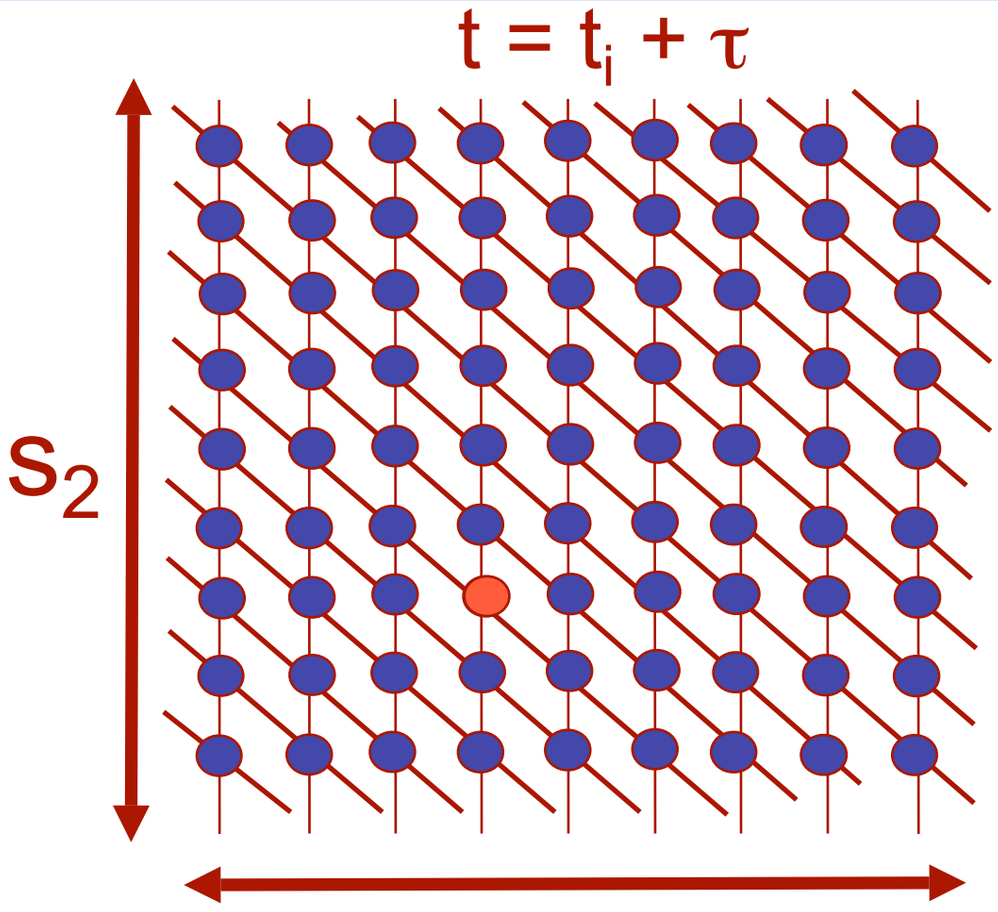
$$\min_{\mu \in \{0,\ldots,M\}} \tau_\mu \in \mathrm{EXP}\,(|\mathbf{w}|_1)$$

The argument, *k*, of this distribution is also a random variable with distribution:

$$P(k = \mu) = \frac{w_\mu}{|\mathbf{w}|_1}$$

In the DM we only generate 2 rv's.

The Direct Method (DM)



Step 1. Generate the time of the next reaction.

The time until the next reaction is a random variable of exponential distribution:

$$\tau \in \mathrm{EXP}\left(|\mathbf{w}|_1\right)$$

To generate the next reaction time, generate r_1 from a uniform distribution on (0,1) and use the equation:

$$\tau = \frac{1}{|\mathbf{w}|_1} \log \frac{1}{r_1}$$

Step 2. Decide which reaction has occurred.

To obtain a realization of which reaction will occur, generate a second uniform random number, r_2, and find the smallest *k* such that:

$$\sum_{\mu=1}^{k-1} w_\mu(\mathbf{x}) \leq r_2 \, |\mathbf{w}|_1 \leq \sum_{\mu=1}^{k} w_\mu(\mathbf{x})$$

Step 3. Update current Time ($t=t+\tau$) and State ($\mathbf{x} = \mathbf{x}+s_k$).

88

The Direct Method
SSA in Matlab.

```
clear all
t=0;tstop = 2000;                                    %%specify initial and final times
x = [0; 0];                                          %% Specify initial conditions
S = [1 -1 0  0; 0  0 1 -1];                          %% Specify stoichiometry
w = inline('[10, 1*x(1), 10*x(1), 1*x(2)]','x');     %% Specify Propensity functions
while t<tstop
    w0 = sum(w(x));                                 % compute the sum of the prop. functions
    t = t+1/w0*log(1/rand);                         % update time of next reaction
    if t<=t_stop
     r2w0=rand*w0;                          % generate second random number and multiply by prop. sum
     i=1;                                           % initialize reaction counter
     while sum(w(1:i))<r2w0                    % increment counter until sum(w(1:i)) exceeds r2w0
       i=i+1;
     end
     x = x+S(:,i);                                  % update the configuration
  end
end
```


89

Limitations on the SSA

- The SSA is an “exact” simulation of the system.

- But…
  - Stepping through every reaction can take a lot of time.
  - A statistical representation of the system dynamics may require many realizations (10^4 to 10^6).

- Faster approximations are available for some problems.

90

Monte-Carlo Simulation Methods

- Stochastic Simulation Algorithm (SSA).
- τ-leaping
  - D. Gillespie, J. Chem. Phys. **115**, 1716 (2001)
  - D. Gillespie, L. Petzold, J. Chem. Phys. **119**, 8229 (2003)
  - M. Rathinam *et al.*, J. Chem. Phys. **119**, 12784 (2003)
  - T. Tian and K. Burrage, J. Chem. Phys. **121**, 10356 (2004)
  - Y. Cao, D. Gillespie and L. Petzold, J. Chem. Phys. **123**, 054104 (2005)

91

τ Leaping

Step 0. Specify length of each time step, τ.

Assume that all propensity functions are constant over the time interval $(t,t+\tau)$.

The number of times each reaction will fire is a Poisson* random number with mean $w_\mu\tau$:

Step 1. For each μ, generate k_μ.

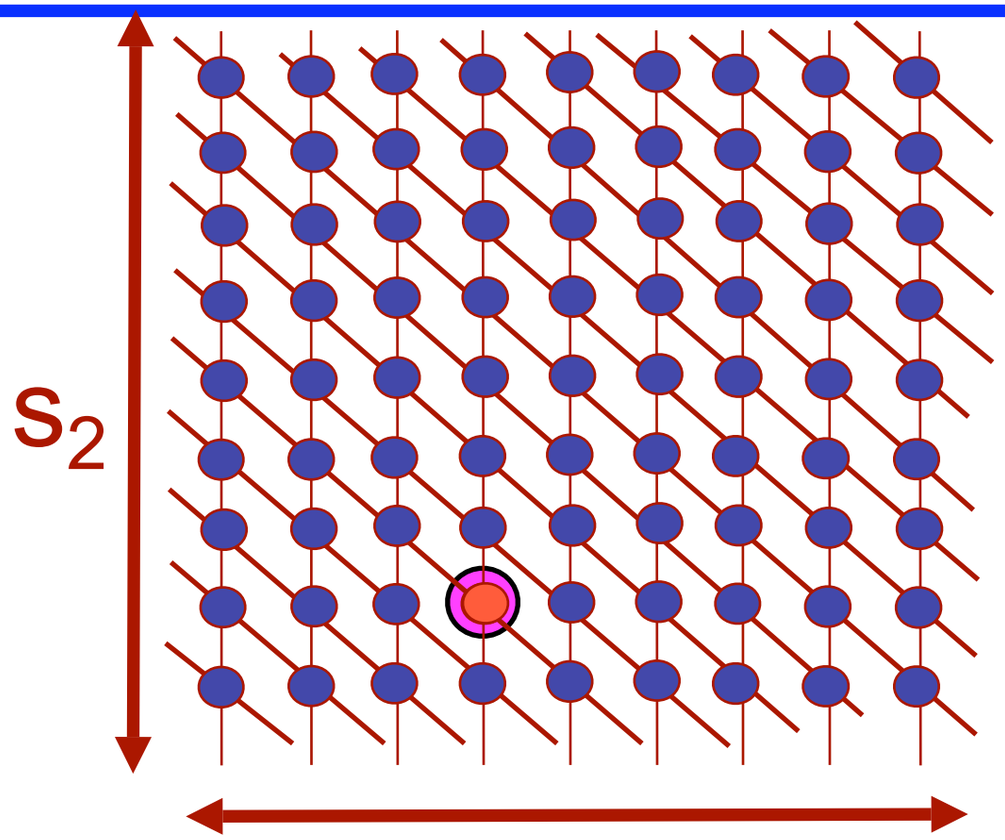
Step 2. Update the time: $t = t + \tau$

Update the state: $\mathbf{x} = \mathbf{x} + \sum_{\mu=1}^{M} k_\mu \mathbf{s}_\mu$

*For some recent studies, binomial RV's are used (T. Tian and K. Burrage, 2004)

92

τ Leaping



t = t_i +τ Update Time

$$k_1 = 4; \mathbf{s}_1 = [0, 1]^T$$

$$k_2 = 2; \mathbf{s}_1 = [-1, 1]^T$$

$$k_3 = 3; \mathbf{s}_1 = [0, -1]^T$$

$$k_4 = 4; \mathbf{s}_1 = [1, -1]^T$$

The number of times each reaction will fire is a Poisson random number with mean $w_\mu \tau$:

Step 1. For each μ, generate k_μ.

Step 2. Update the state: $\mathbf{x} = \mathbf{x} + \sum_{\mu=1}^{M} k_\mu \mathbf{s}_\mu$

Update the time: $t = t + \tau$

93

Limitations of τ leaping

- For many situations τ leaping significantly speeds up the Monte Carlo simulation, but:
  - Poisson r.v.'s are unbounded
    - Propensity functions may change dramatically over small time intervals.
  - May result in negative populations.

Note that these concerns are most important when the population of some species are very small.

Precisely the circumstance where stochastic models are most important!

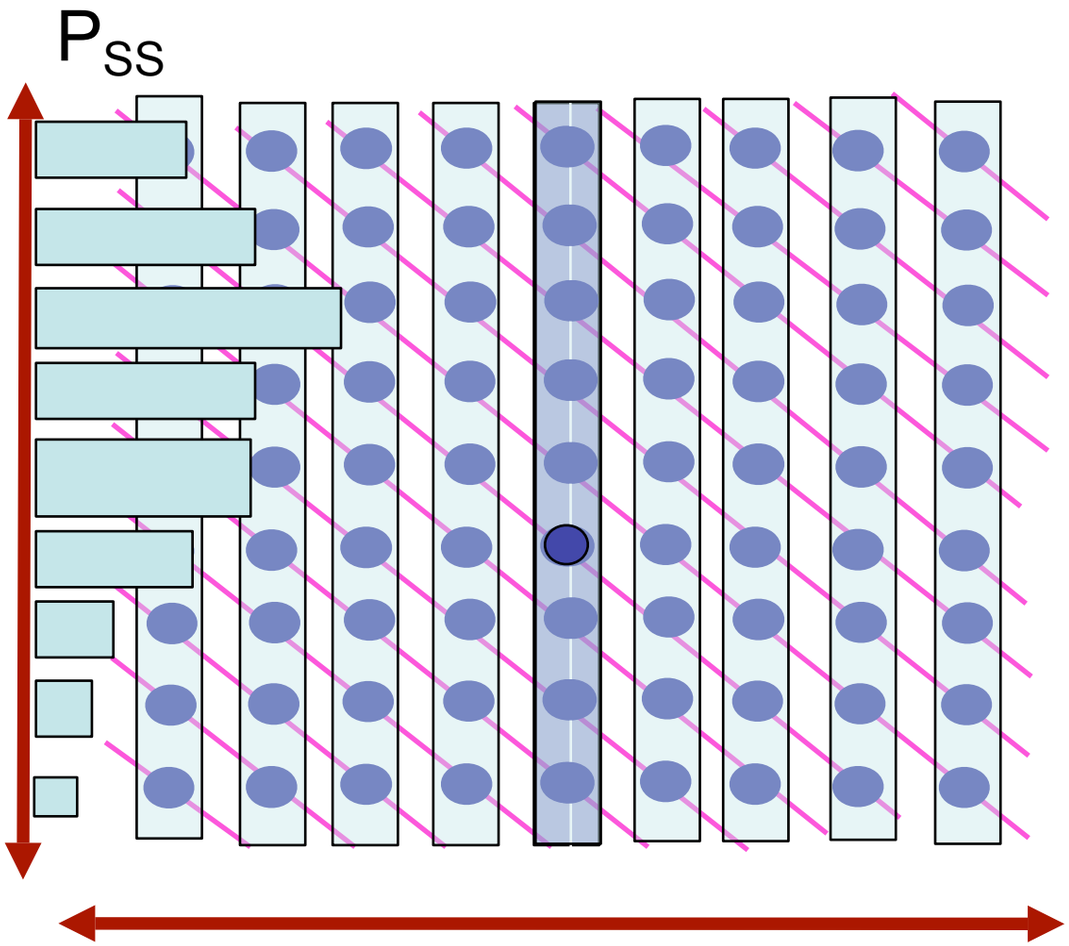
Chemical Langevin Equation

95

Monte-Carlo Simulation Methods

- Stochastic Simulation Algorithm (SSA).
- τ-leaping
- System Partitioning Methods
  - Fast--Slow Partitions
    - C. Rao and A. Arkin, J. Chem. Phys. **118**, 4999 (2003)
    - Y. Cao *et al.*, J. Chem. Phys. **122**, 014116 (2005)
  - Continuous--Discrete Partitions
    - E. Haseltine and J. Rawlings, J. Chem. Phys. **117**, 6959 (2002)
    - H. Salis and Y. Kaznessis, J. Chem. Phys. **122**, 054103 (2005)

Fast--Slow partitions.



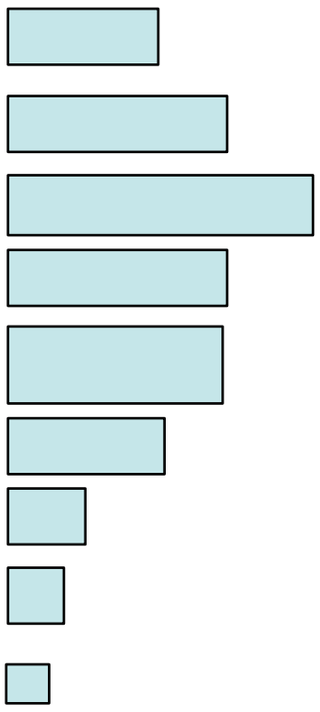
Separate into “fast” and “slow” partitions.

Assume that the “fast” partitions reach probabilistic equilibrium before a slow reaction occurs.

97

Fast--Slow partitions.

P_{SS}



Slow Reaction Propensities

$$\times \begin{bmatrix} w_\mu(\mathbf{x}_1) \\ w_\mu(\mathbf{x}_2) \\ w_\mu(\mathbf{x}_3) \\ \vdots \end{bmatrix}$$

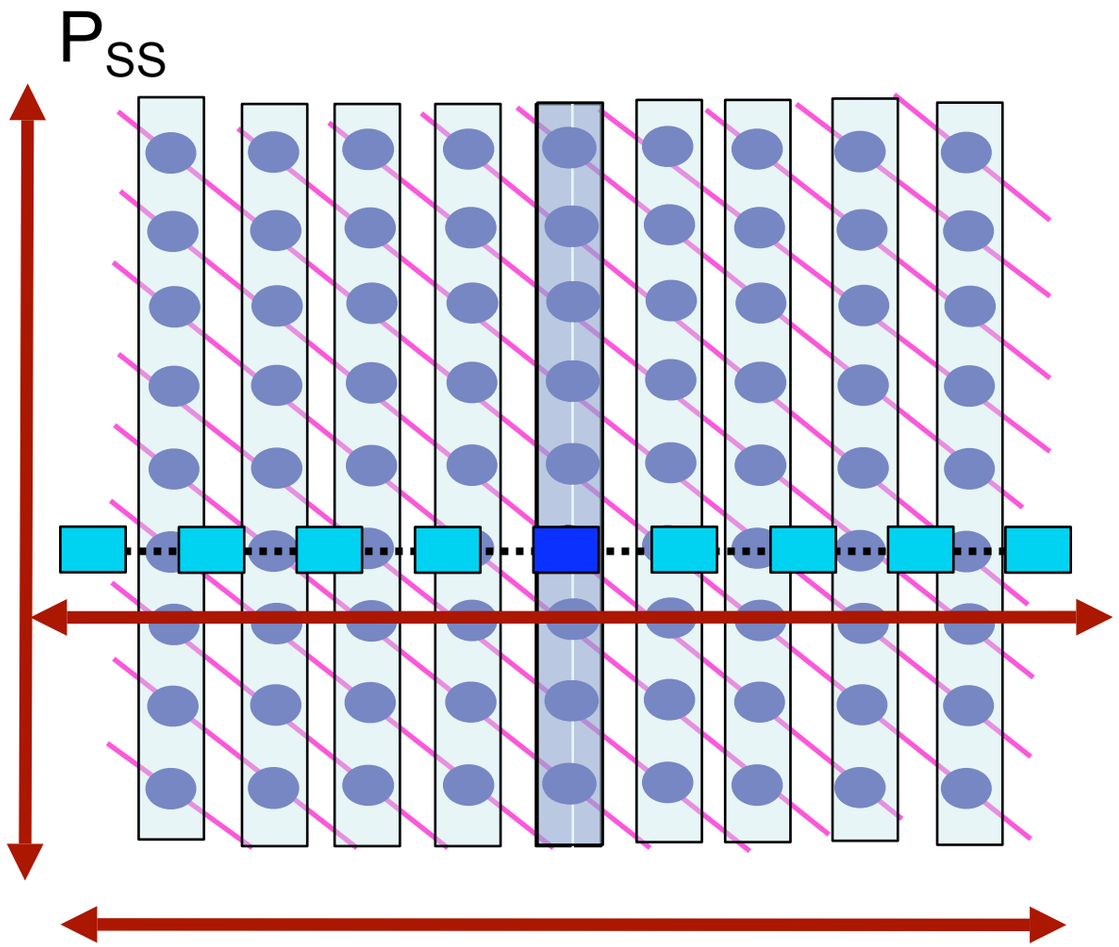
Average Slow Reaction Propensities

$$= \bar{w}_\mu, \text{ for } \mu = \{1, 2, \ldots, M\}$$

Use the fast sets' steady state probability distributions to scale the propensity functions of the slow reactions.

Results in a vector of average propensity functions, $\bar{\mathbf{w}}$, for the slow reactions.

Fast--Slow partitions.



The projection to the slow manifold results in a new lower dimensional Markov chain.

This is simulated with SSA.

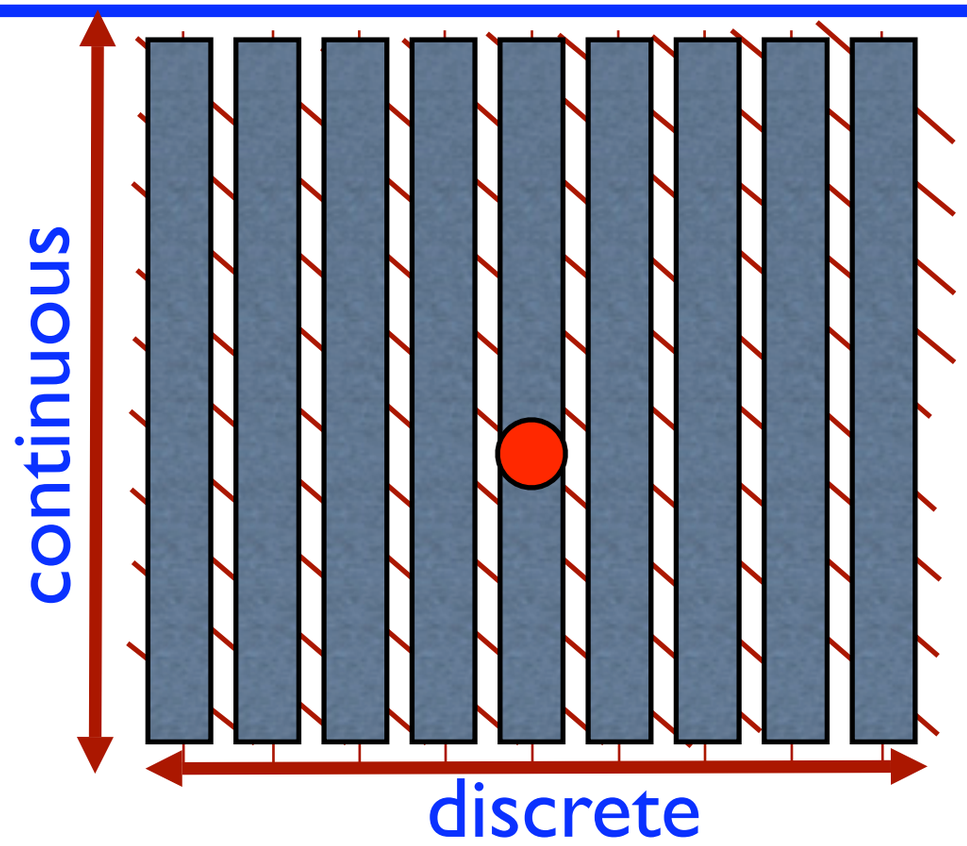
99

Continuous--Discrete partitions.

- In some systems, there are great differences in scale:
  - Large populations (continuous)
  - Small populations (discrete)
- All discrete models take too long.
- All continuous models are inaccurate.
- Hybrid models are necessary.

100

Separate into "continuous" and "discrete" partitions.

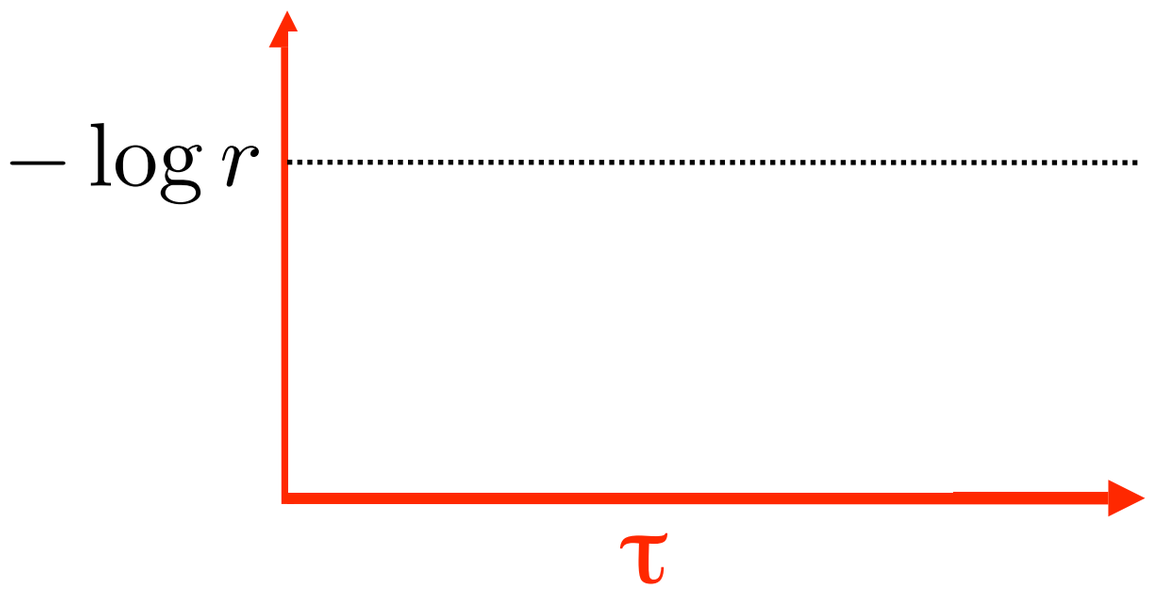


Simulate the continuous part with ordinary or stochastic differential equations.

Choose uniform rv, *r*.

Numerically integrate propensity functions until:

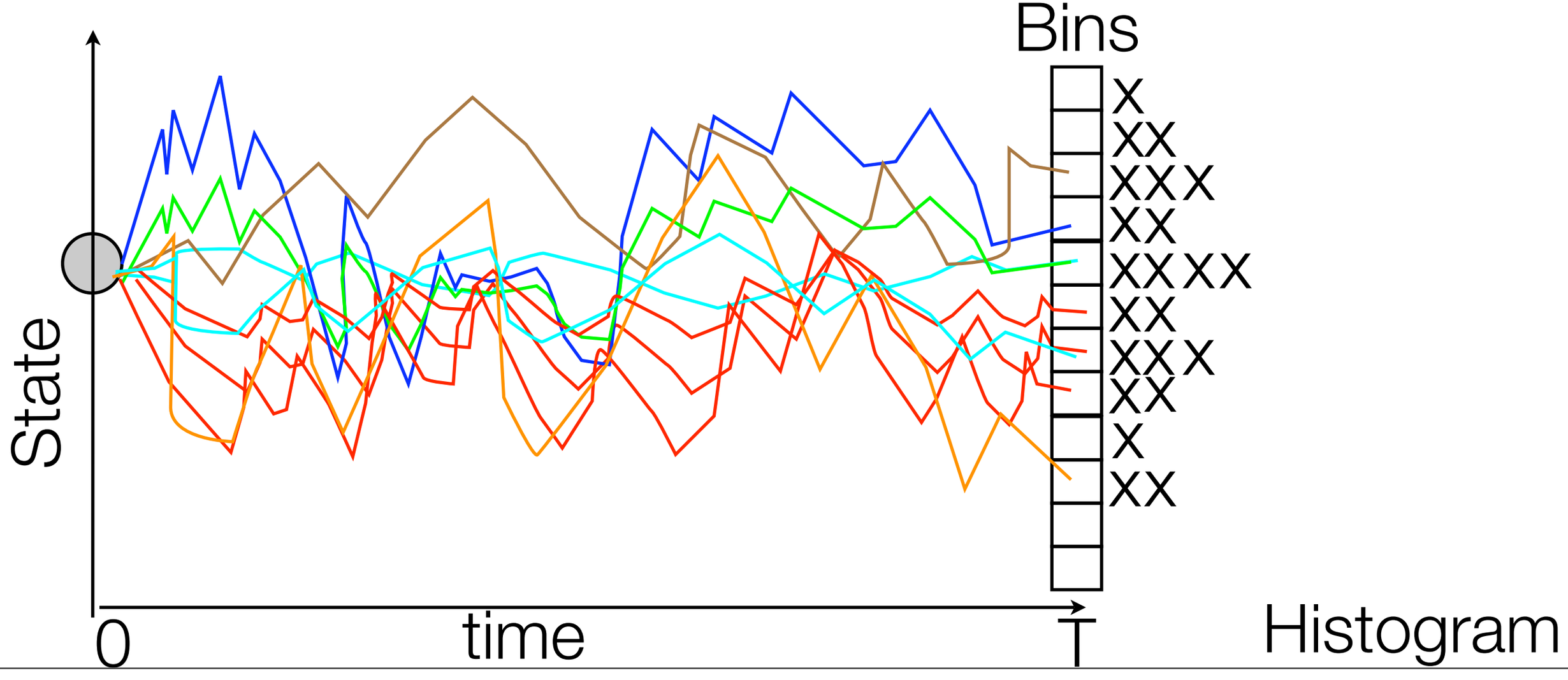
$$\int_{t_0}^{t_0+\tau} \sum_{\mu=1}^{M} w_\mu(\mathbf{x}(t))dt = -\log r$$



Choose next discrete reaction.

Using the SSA to Find Distributions

- The SSA does an excellent job of producing possible trajectories.



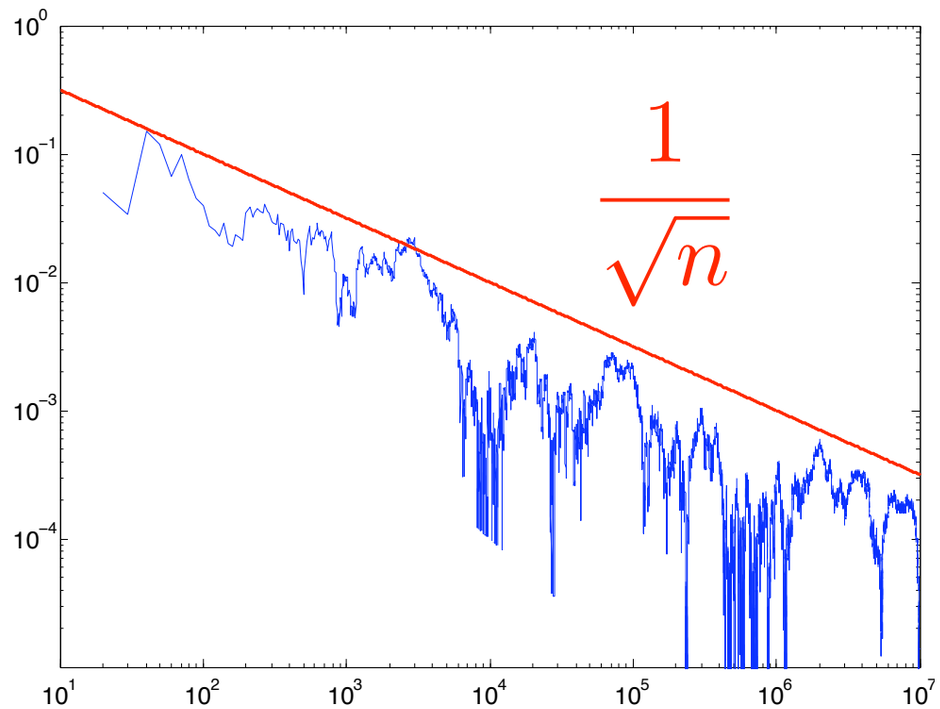
Convergence of the SSA

- To get more accurate distributions, one needs more SSA runs.
- Unfortunately, the convergence rate of any Monte Carlo algorithm is fundamentally limited: $error = \mathcal{O}(n^{-\frac{1}{2}})$
- If very high precision is required, then MC methods will be very inefficient.

Convergence for Coin Toss

error:

$$\left| \frac{Heads}{n} - 0.5 \right|$$



After 10^7 tosses there is still an error of about 3×10^{-4}.

103

☑ Monte Carlo Solution Schemes

104

Density Computations

- The Finite State Projection (FSP) solution to the Chemical Master Equation.

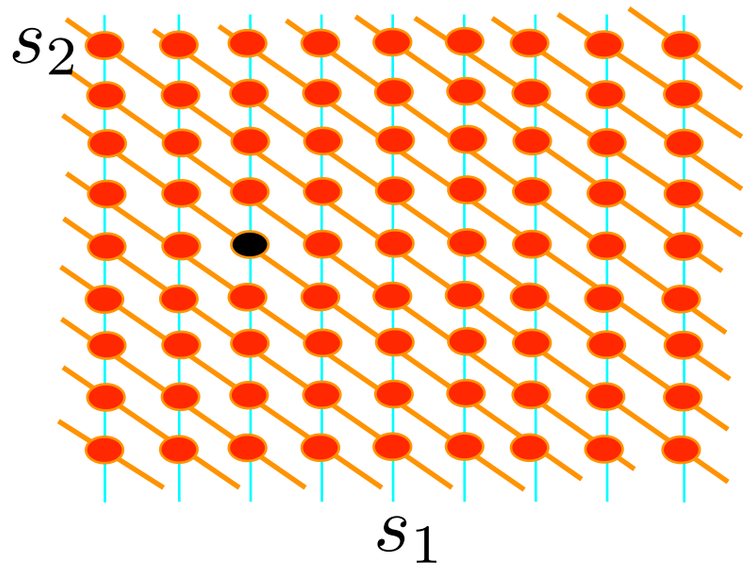
- Reductions to the FSP

- Case studies.

106

The Chemical Master Equation

The probability that the system is in configuration **x** at *t+dt* is equal to the probability that the system is at **x** at *t*, and no reaction occurs between *t* and *t+dt* plus the probability that the system is one reaction removed from **x** at *t* and that reaction occurs between *t* and *t+dt*.



The CME (McQuarrie '67):

$$\dot{p}(\mathbf{x},t) = -p(\mathbf{x},t)\sum_{k=1}^{M} w_k(\mathbf{x}) + \sum_{k=1}^{M} p(\mathbf{x}-\mathbf{s}_k,t)w_k(\mathbf{x}-\mathbf{s}_k)$$

Define the probability density state vector (pdv): $\mathbf{P}(\mathbf{X},t) := [p(\mathbf{x}_1,t), p(\mathbf{x}_2,t), p(\mathbf{x}_3,t), \ldots]^T$.

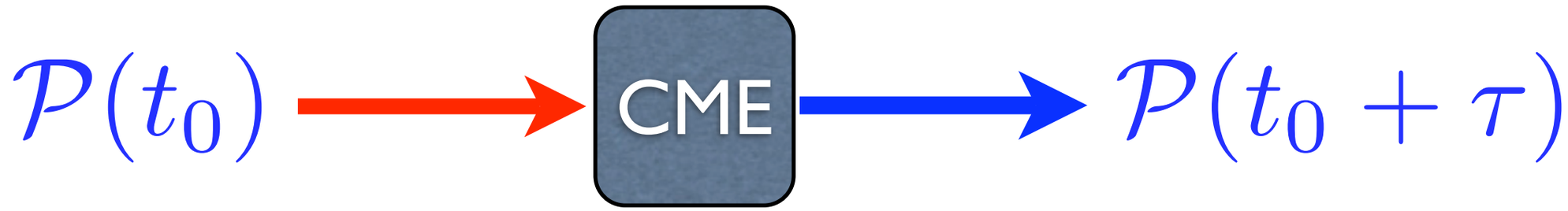
$\mathbf{P}(\mathbf{X},t)$ evolves according to the Linear Time Invariant ODE:

$$\dot{\mathbf{P}}(\mathbf{X},t) = \mathbf{A}\cdot\mathbf{P}(\mathbf{X},t)$$

The matrix CME

The Chemical Master Equation

- The solution of the CME is a transfer operator:



- The dimension of the CME can be INFINITE.
  - Most CME's cannot be solved, so approximations are needed.

108

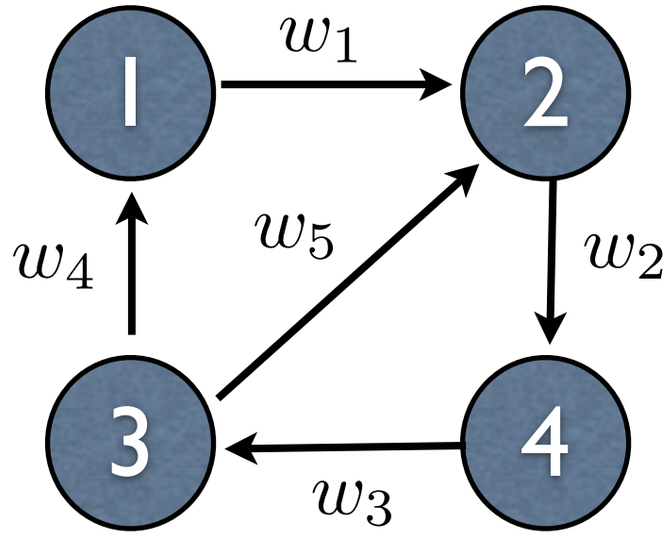
Forming the Generator

A has one row/column for each state.

Each transition, $\mathbf{x}_i \rightarrow \mathbf{x}_j$, contributes to **A** in two locations:

$-w_\mu(\mathbf{x}_i)$ goes in the *diagonal* element $A_{i,i}$

$+w_\mu(\mathbf{x}_i)$ goes in the *off-diagonal* element $A_{j,i}$



$$\mathbf{A} = \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix}$$

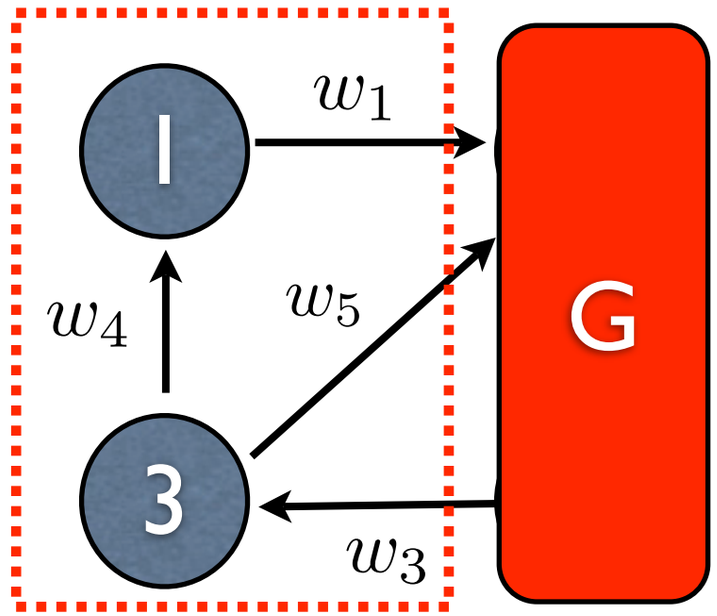
109

The Finite State Projection

Select the states to keep.

Find the corresponding projection matrix:

$$\mathbf{A}_{[1,3]} = \begin{bmatrix} -w_1 & w_4 \\ 0 & -w_4 - w_5 \end{bmatrix}$$



$$\mathbf{A} = \begin{bmatrix} -w_1 & 0 & w_4 & 0 \\ w_1 & -w_2 & w_5 & 0 \\ 0 & 0 & -w_4 - w_5 & w_3 \\ 0 & w_2 & 0 & -w_3 \end{bmatrix}$$

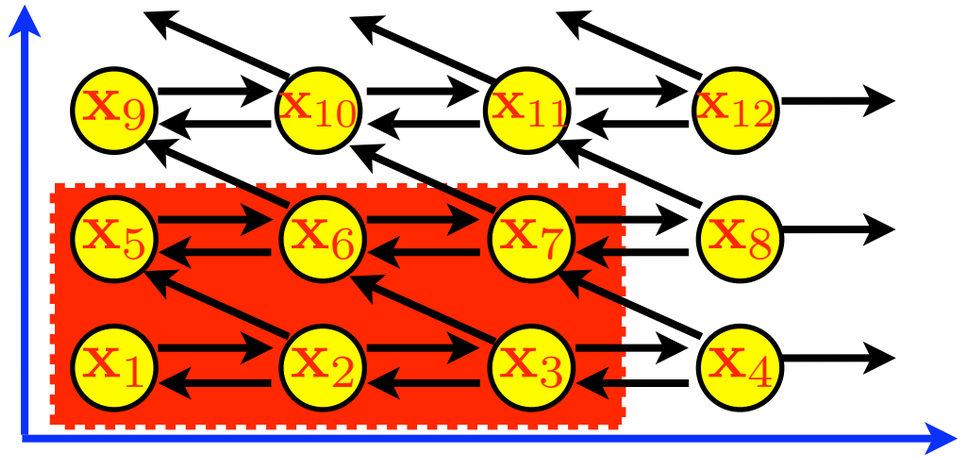
Collapse remaining states into a single absorbing state

$$\mathbf{A}^{FSP}_{[1,3]} = \begin{bmatrix} -w_1 & w_4 & 0 \\ 0 & -w_4 - w_5 & 0 \\ w_1 & w_5 & 0 \end{bmatrix}$$

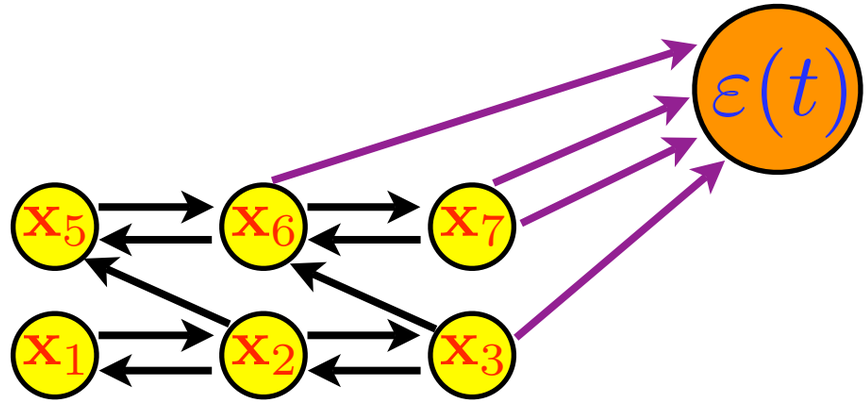
This is the generator for a new Markov chain

The Finite State Projection Method

The Full System



The Projected System (FSP)



Full Master Equation

$$\begin{bmatrix} \dot{\mathbf{P}}_J \\ \dot{\mathbf{P}}_{J'} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_J & \mathbf{A}_{JJ'} \\ \mathbf{A}_{J'J} & \mathbf{A}_{J'} \end{bmatrix} \begin{bmatrix} \mathbf{P}_J(t) \\ \mathbf{P}_{J'}(t) \end{bmatrix}$$

Dimension = $\#(J) + \#(J')$ = Infinite

FSP Master Equation

$$\begin{bmatrix} \dot{\mathbf{P}}_J^{FSP} \\ \dot{\varepsilon} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_J & \mathbf{0} \\ -\mathbf{1}^T\mathbf{A}_J & 0 \end{bmatrix} \begin{bmatrix} \mathbf{P}_J^{FSP}(t) \\ \varepsilon(t) \end{bmatrix}$$

Dimension = $\#(J)+1$ = **7**

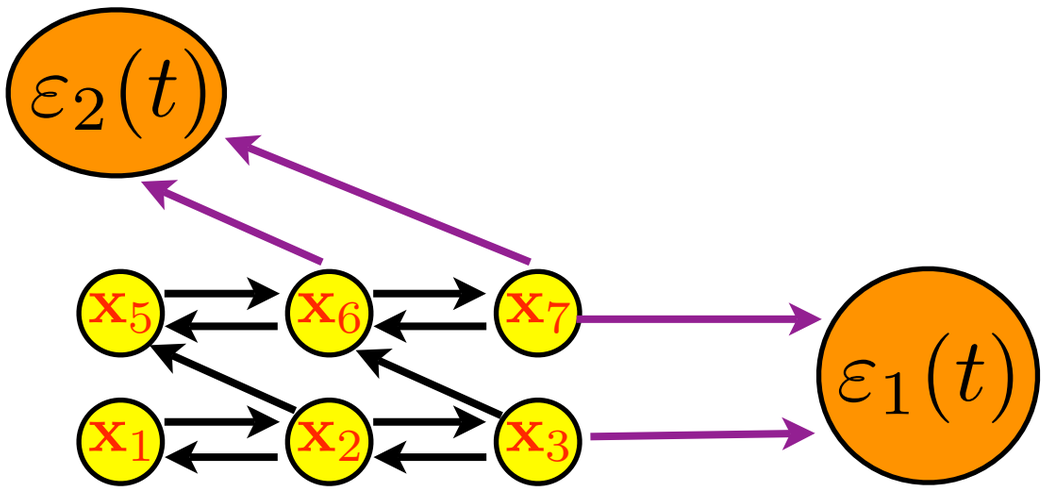
The FSP Theorem

(Munsky/Khammash JCP '06)

$$\mathbf{P}_J(t) \geq \mathbf{P}_J^{FSP}(t) \text{ and}$$

$$\left\| \begin{bmatrix} \mathbf{P}_J(t) \\ \mathbf{P}_{J'} \end{bmatrix} - \begin{bmatrix} \mathbf{P}_J^{FSP}(t) \\ \mathbf{0} \end{bmatrix} \right\|_1 = \varepsilon(t)$$

A Test...



What do $\varepsilon_1(t)$ and $\varepsilon_2(t)$ mean?

112

The Finite State Projection Algorithm

Inputs:
Initial Conditions, System Parameters,
Final time (t_f), Allowable error ($\varepsilon_{\max}$)

↓

Step 1:
Choose initial projection space, $\mathbf{X}_{J_0}$.

↓

Step 2:
Use projection $\mathbf{X}_{J_i}$ to find corresponding error, $\varepsilon_i(t_f)$.

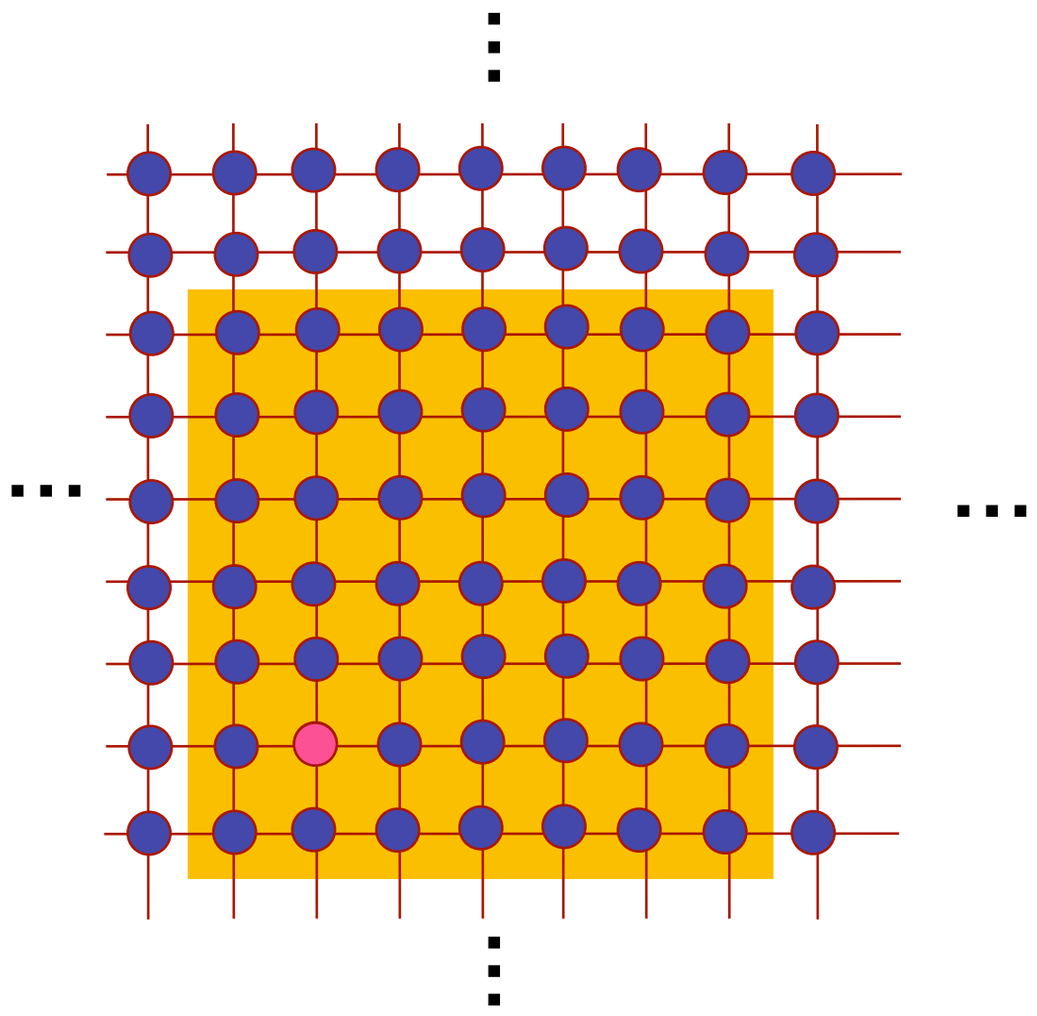
↓

Step 3:
If $\varepsilon_i(t_f) \leq \varepsilon_{\max}$, **Stop**.
$\mathcal{P}_{J_i}^{FSP}(t_f)$ approximates $\mathcal{P}(t_f)$ to within $\varepsilon_{\max}$.

↓

Step 4:
Expand projection, $\mathbf{X}_{J_{i+1}} \supset \mathbf{X}_{J_i}$,
Increment i and return to Step 2.

The FSP Algorithm



Begin with initial conditions, process parameters, and error tolerance.
Choose an initial set: $\mathbf{X}_{J_0}$.

Find ε_i;

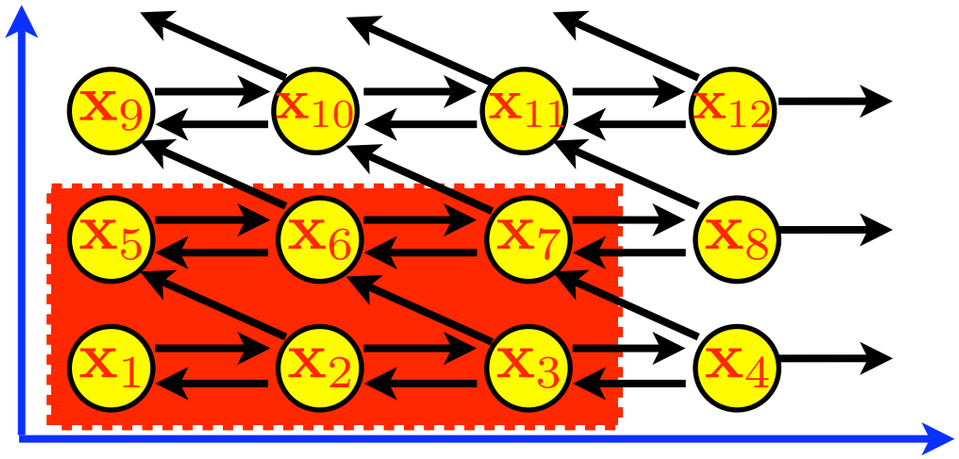
If $\varepsilon_i < \varepsilon_{max}$, STOP.

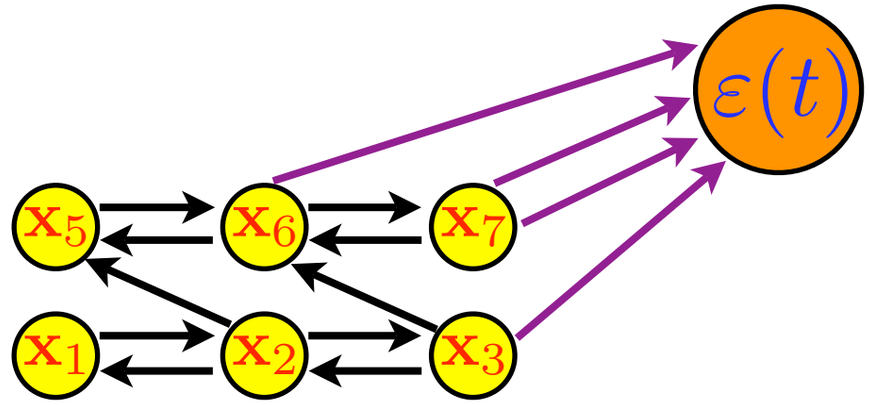
Otherwise add more configurations to get $\mathbf{X}_{J_{i+1}}$.

$\varepsilon_0 > \varepsilon_{max}$ $\varepsilon_2 > \varepsilon_{max}$ $\varepsilon_4 > \varepsilon_{max}$

$\varepsilon_1 > \varepsilon_{max}$ $\varepsilon_3 > \varepsilon_{max}$ $\varepsilon_5 < \varepsilon_{max}$ STOP

The "error" sink of the FSP to get exit times.

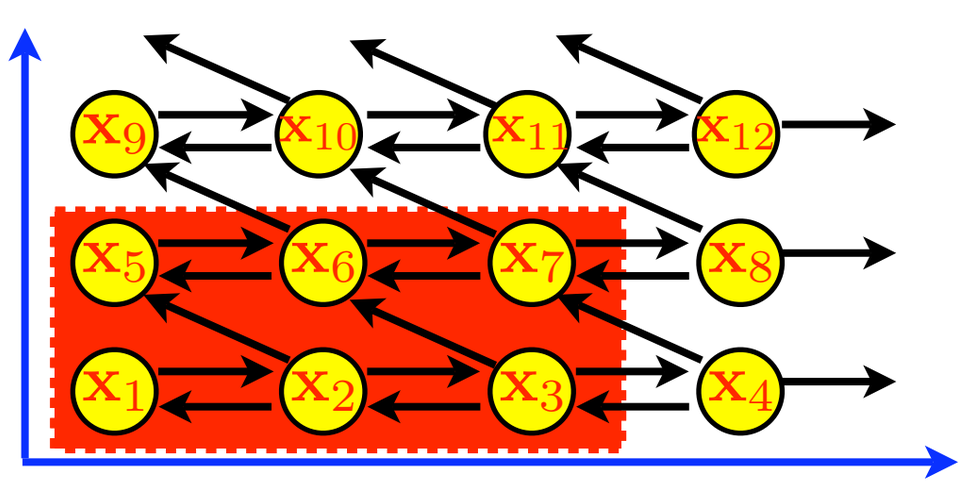


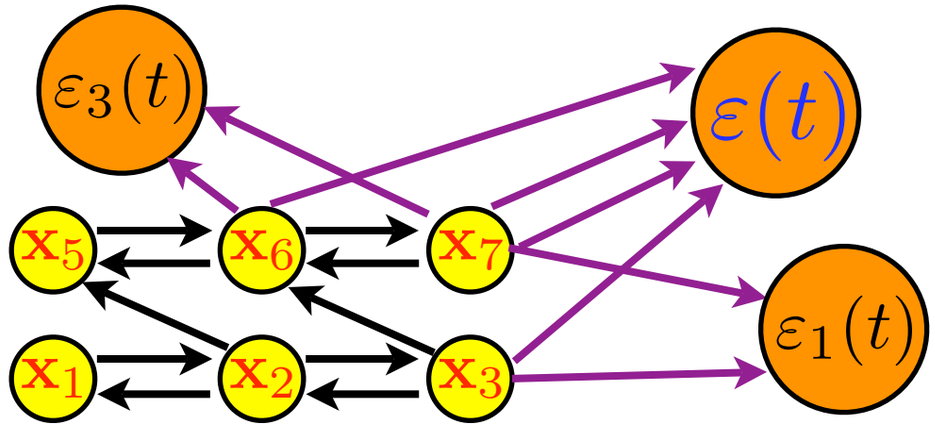


- In the original FSP, $\varepsilon(t)$ is the amount of the probability measure that exits the projection region $\mathbf{X}_J$.
- Median exit time: $t_{50} = t,\ \text{s.t.}\ \varepsilon(t) = 0.5$
- In this form $\varepsilon(t)$ gives information as to when the system exits $\mathbf{X}_J$, but not how.

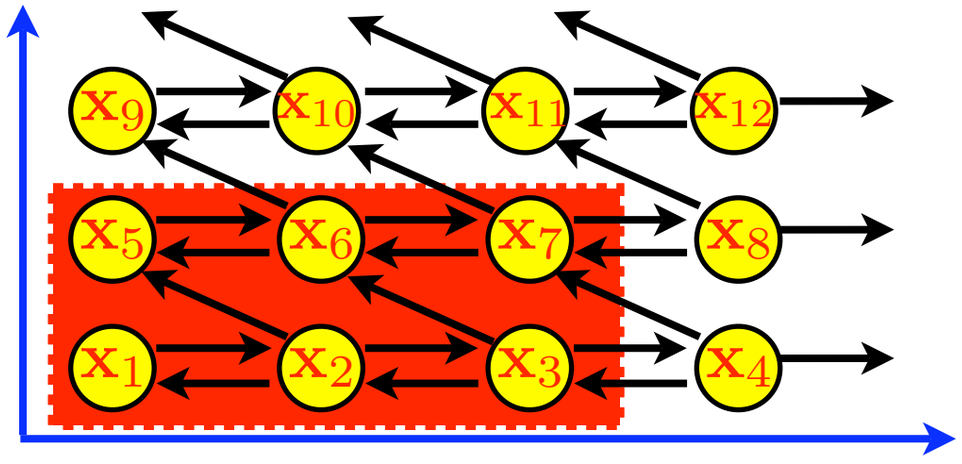
Multiple FSP sinks to get exit directions.

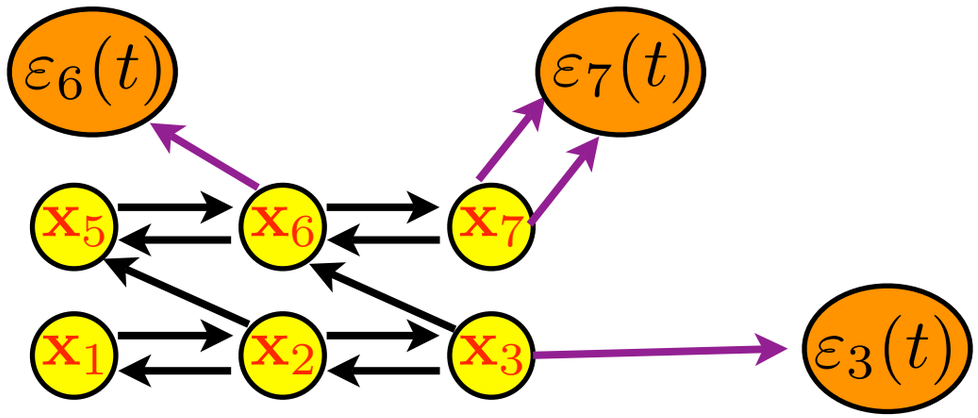
- By using multiple sinks, one can determine *how* the probability measure exits $\mathbf{X}_J$.





Which Reaction Leaves $\mathbf{X}_J$?





From which state?

116

Advantages of the FSP.

- Deterministic.
  - ★ Every run of the FSP yields the same result.
  - ★ Enables easier comparisons of different systems (sensitivity analysis).
- Provides accuracy guarantees.
  - ★ Can be made as precise as required.
  - ★ Allows for analysis of rare events.
- Does not depend upon initial conditions.
- Is open to many subsequent model reductions.

117

Limitations

- Numerical stiffness may lead to computational inefficiency.
- Systems may become very large as distributions cover large regions of the configuration space.
  - ★ Compact distributions may drift over time.
  - ★ Dilute distributions may spread over large regions.
  - ★ Dimension grows exponentially with the number of species.
- For these problems, the original FSP may not suffice,
- BUT, with additional model reductions and systematic techniques, many of these problems may be alleviated.

118

Outline

- ☑ Finite State Projection (FSP)
- Reductions to the FSP
  - ★ Aggregating unobservable states
    - Munsky/Khammash, *CDC*, 2006
  - ★ Time interval discretization
  - ★ Slow manifold projection
  - ★ Coarse meshes for the CME

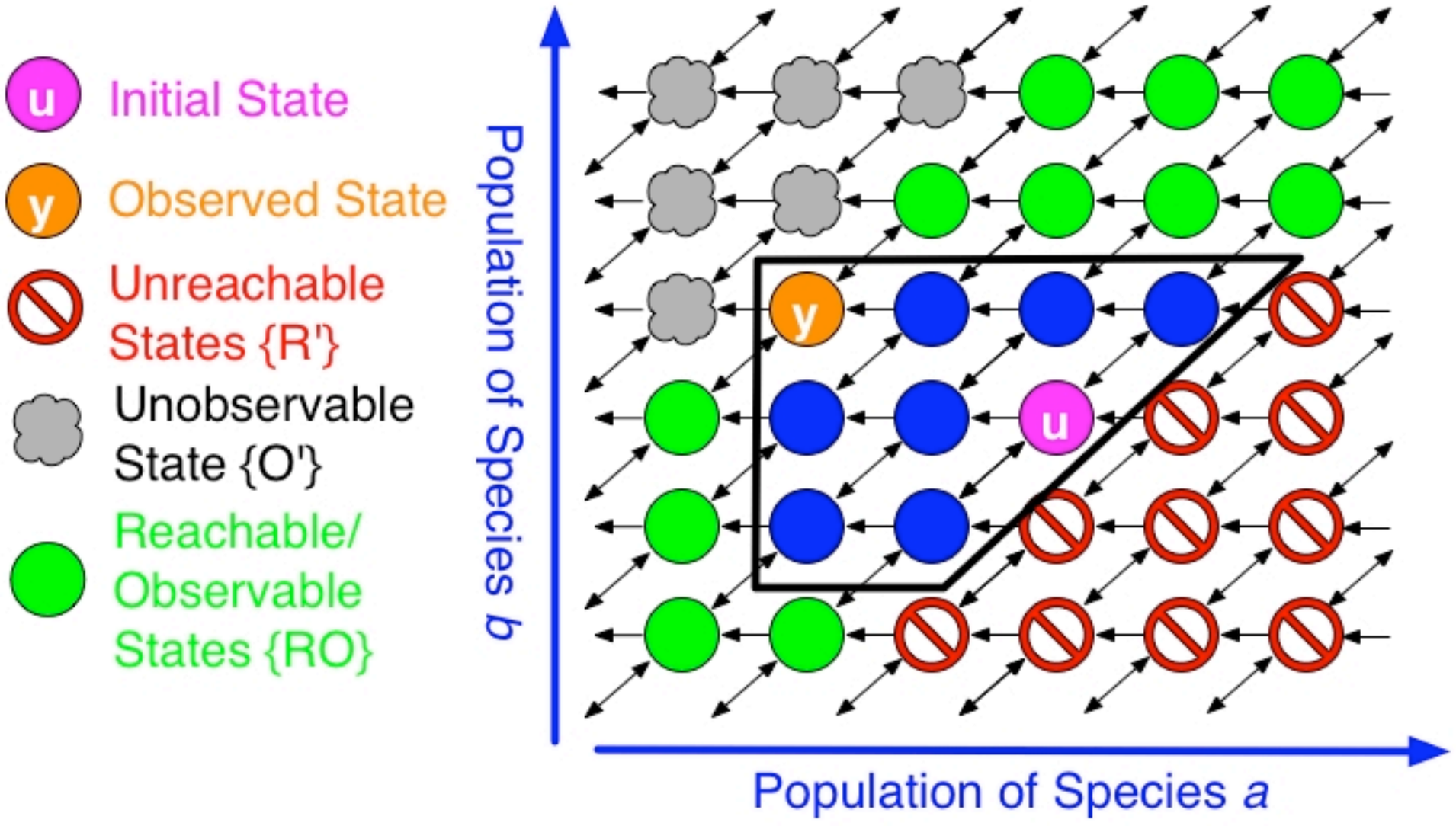
119

Using Input & Output relations for model reduction.

- Often one is not interested in the entire probability distribution.
- Instead one may wish only to estimate:
  - ★ a statistical summary of the distribution, e.g.
    - ✦ means, variances, or higher moments
  - ★ probability of certain traits:
    - ✦ switch rate, extinction, specific trajectories, etc…
- In each of these cases, one can define an output **y**(t):

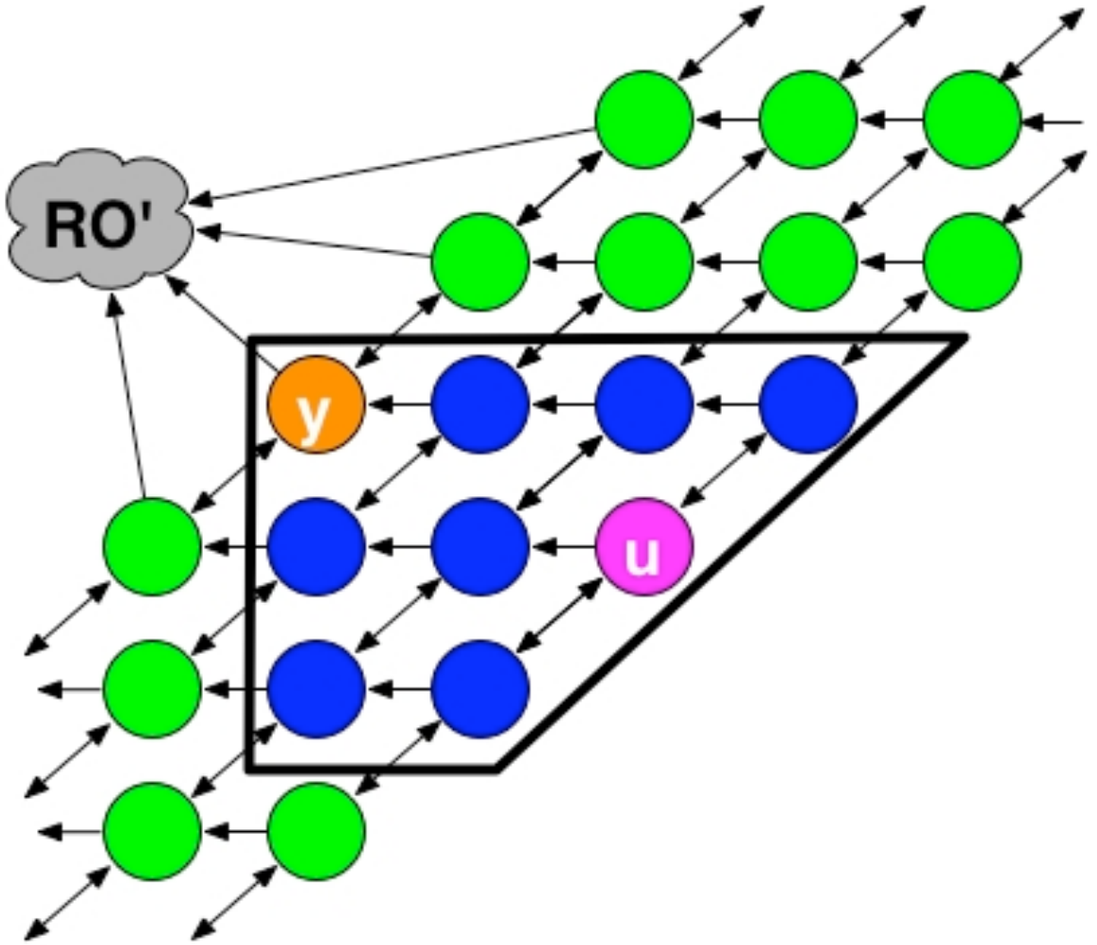
$$\dot{\mathbf{P}}(t) = \mathbf{A}\mathbf{P}(t)$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{P}(t)$$

Begin with a Full Integer Lattice Description of the System States.



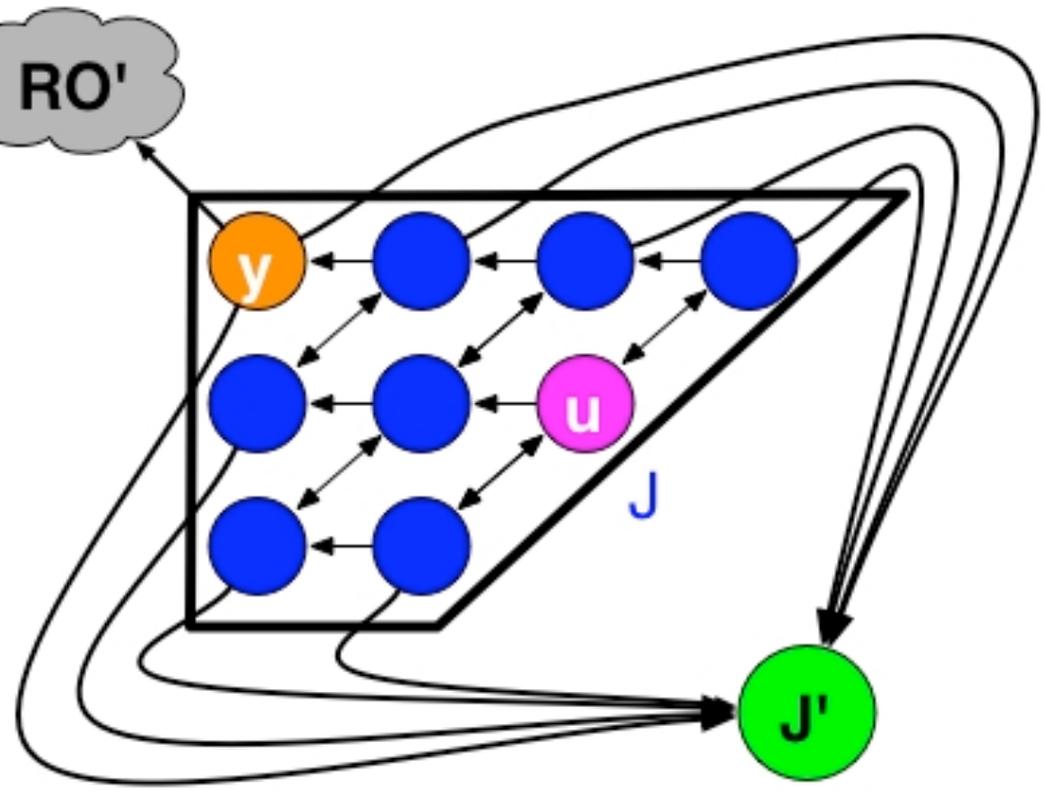
121

Remove Unreachable States and Aggregate the Observable States.



122

Project the Reachable/Observable States onto a Finite Subspace.



We now have a solvable approximation, for which the FSP gives bounds on the approximation's accuracy.

123

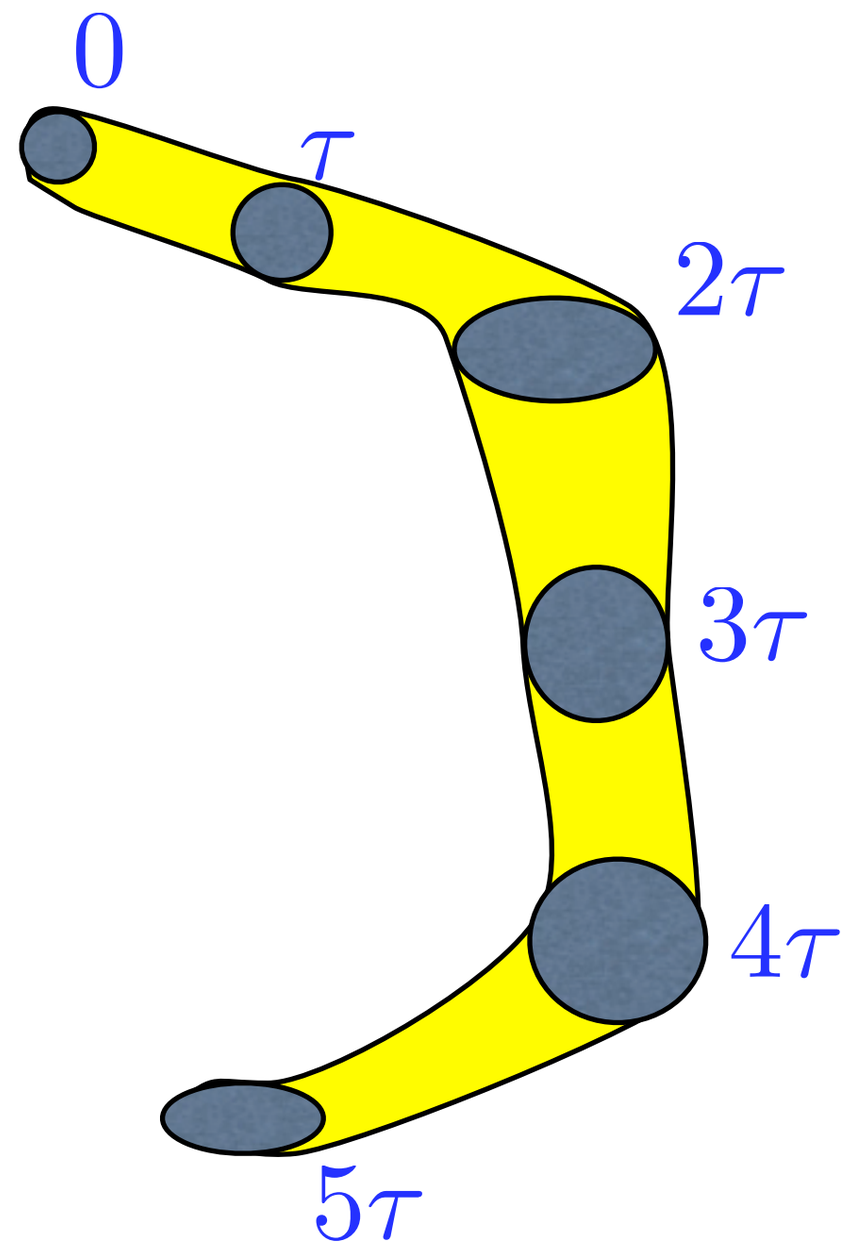
Outline

- ☑ Finite State Projection (FSP)
- Reductions to the FSP
  - ★ Aggregating unobservable states
  - ★ Time interval discretization
    - Munsky and Khammash, *J. Comp. Phys.*, 2007
    - Burrage et al, *A.A. Markov 150th Anniv. Meeting*, 2006
  - ★ Slow manifold projection
  - ★ Coarse meshes for the CME

124

Time Interval Discretization for the FSP

- ★ For many systems, the distribution may drift over time.
- ★ At any one time, the distribution may have a limited support, but...
- ★ The FSP solution must include all intermediate configurations.
- ★ This may lead to an exorbitantly large system of ODEs.

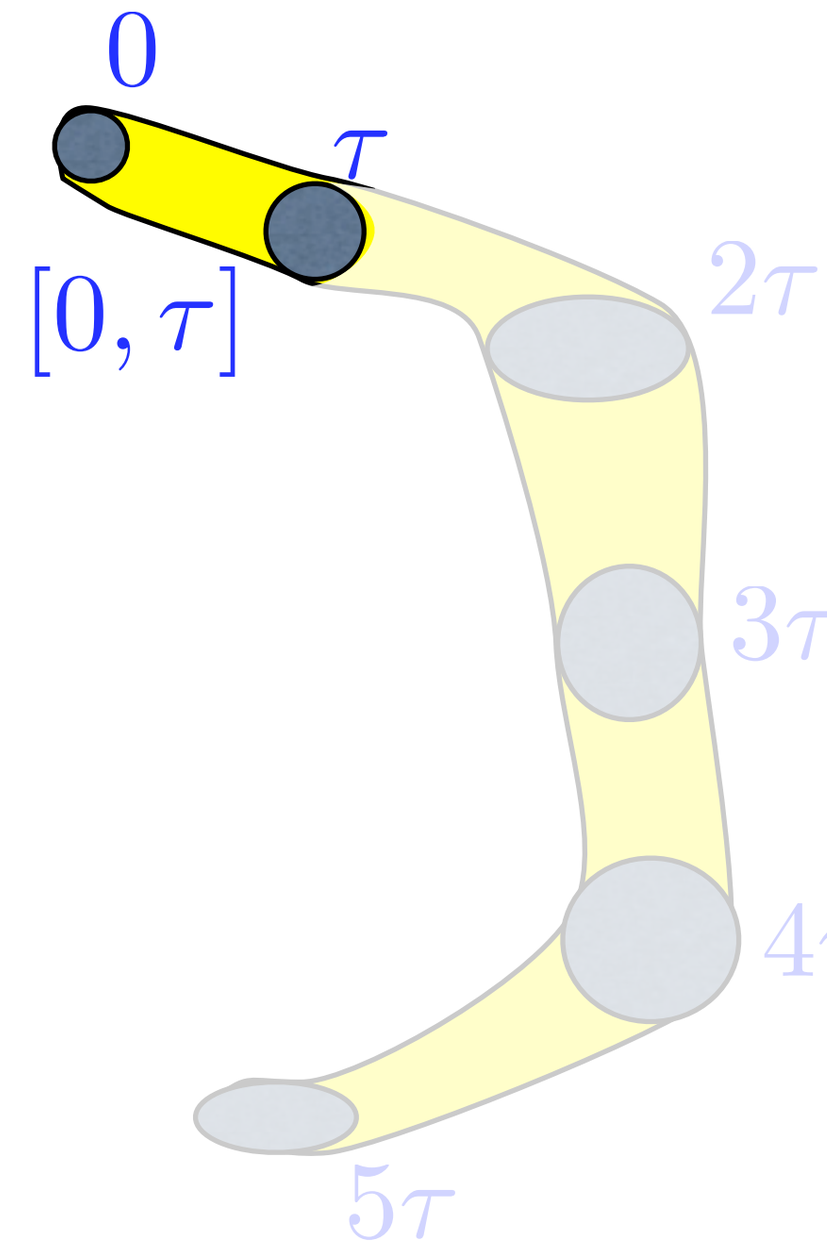


125

Time Interval Discretization for the FSP

★ Instead:

✳ Discretize the time interval into smaller steps and solve a separate projection for each interval.

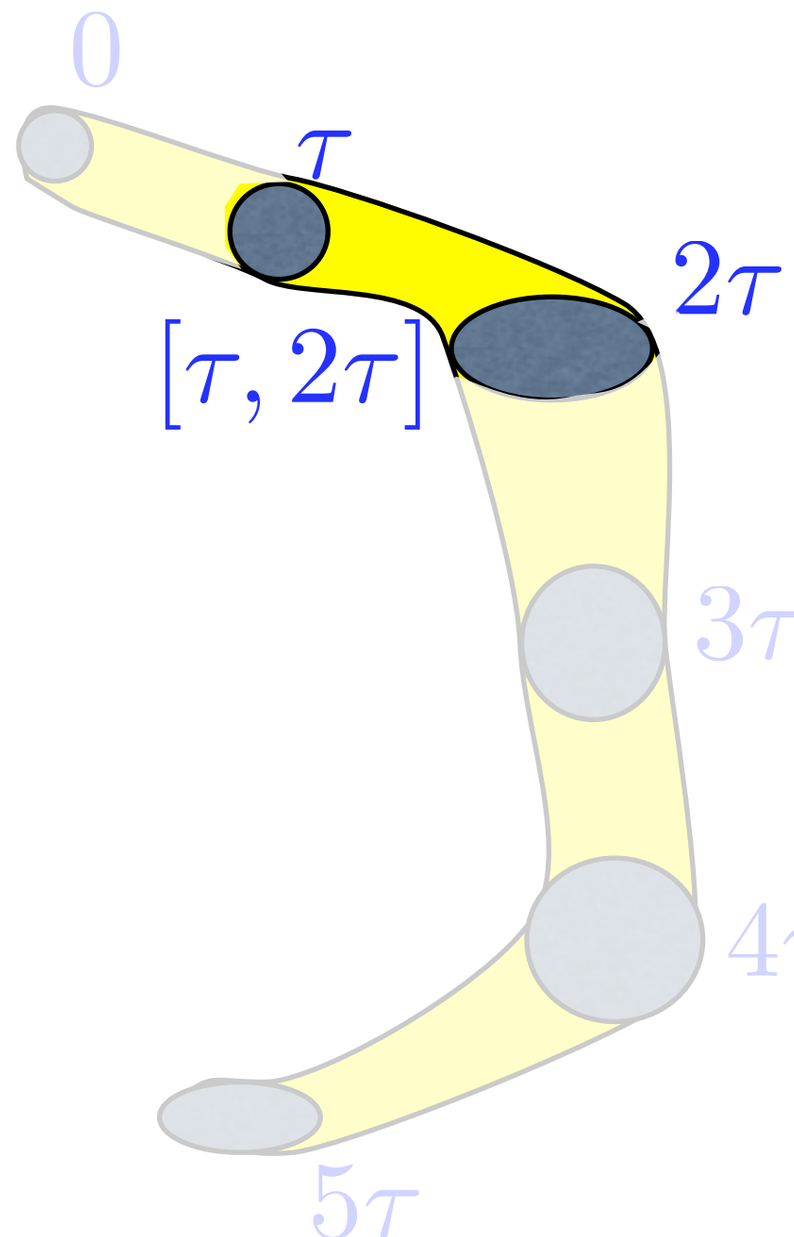


126

Time Interval Discretization for the FSP

★ Instead:

✳ Discretize the time interval into smaller steps and solve a separate projection for each interval.

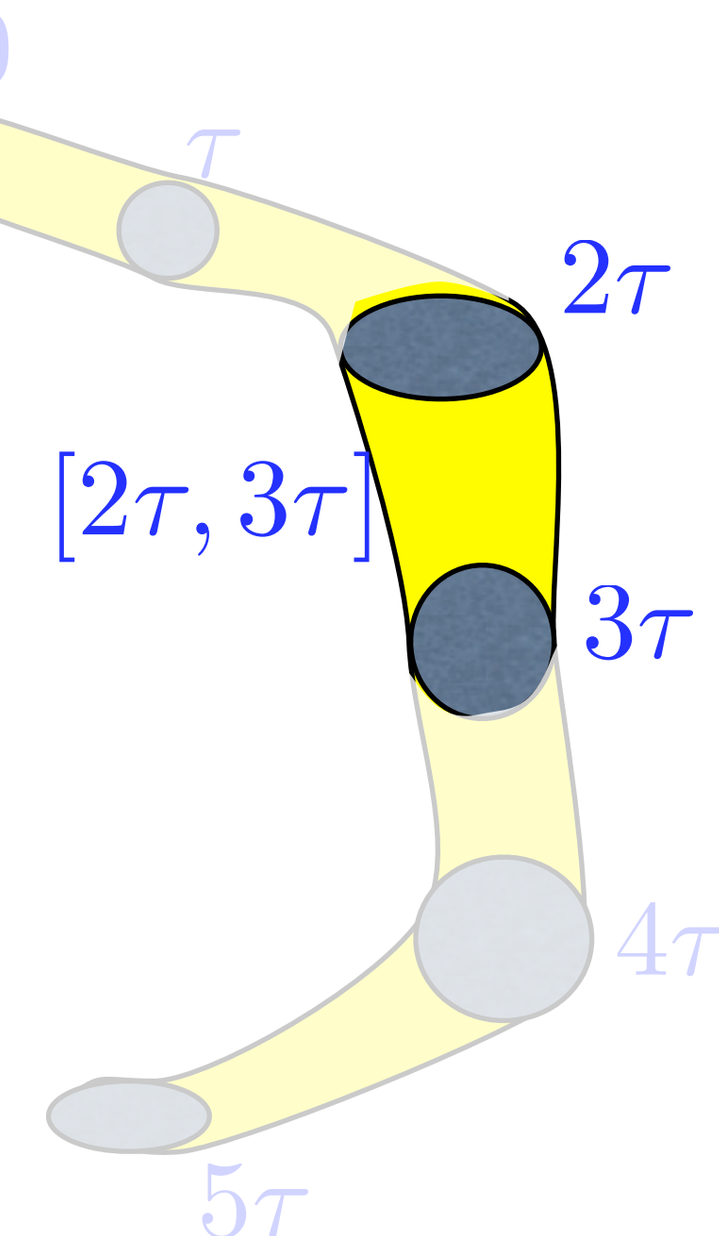


127

Time Interval Discretization for the FSP

★ Instead:

✳ Discretize the time interval into smaller steps and solve a separate projection for each interval.

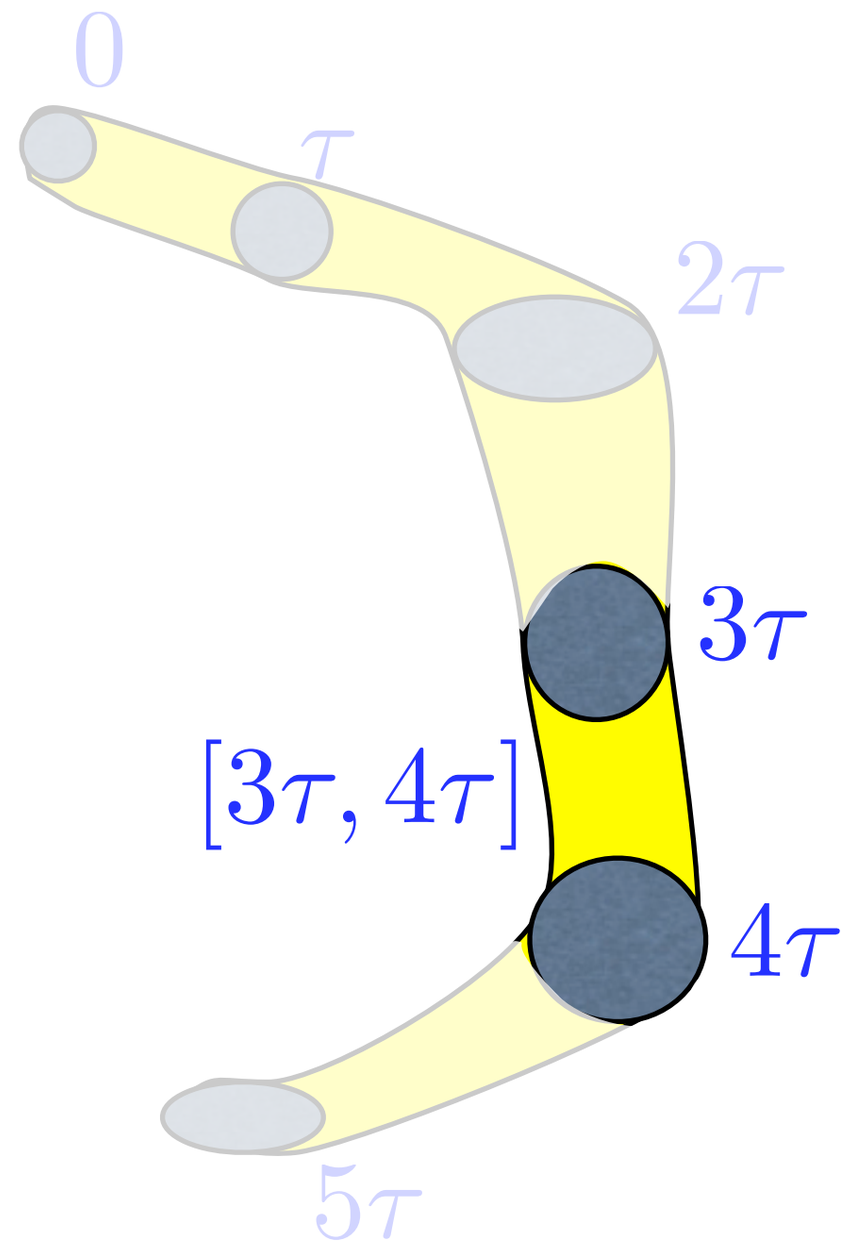


128

Time Interval Discretization for the FSP

★ Instead:

✴ Discretize the time interval into smaller steps and solve a separate projection for each interval.

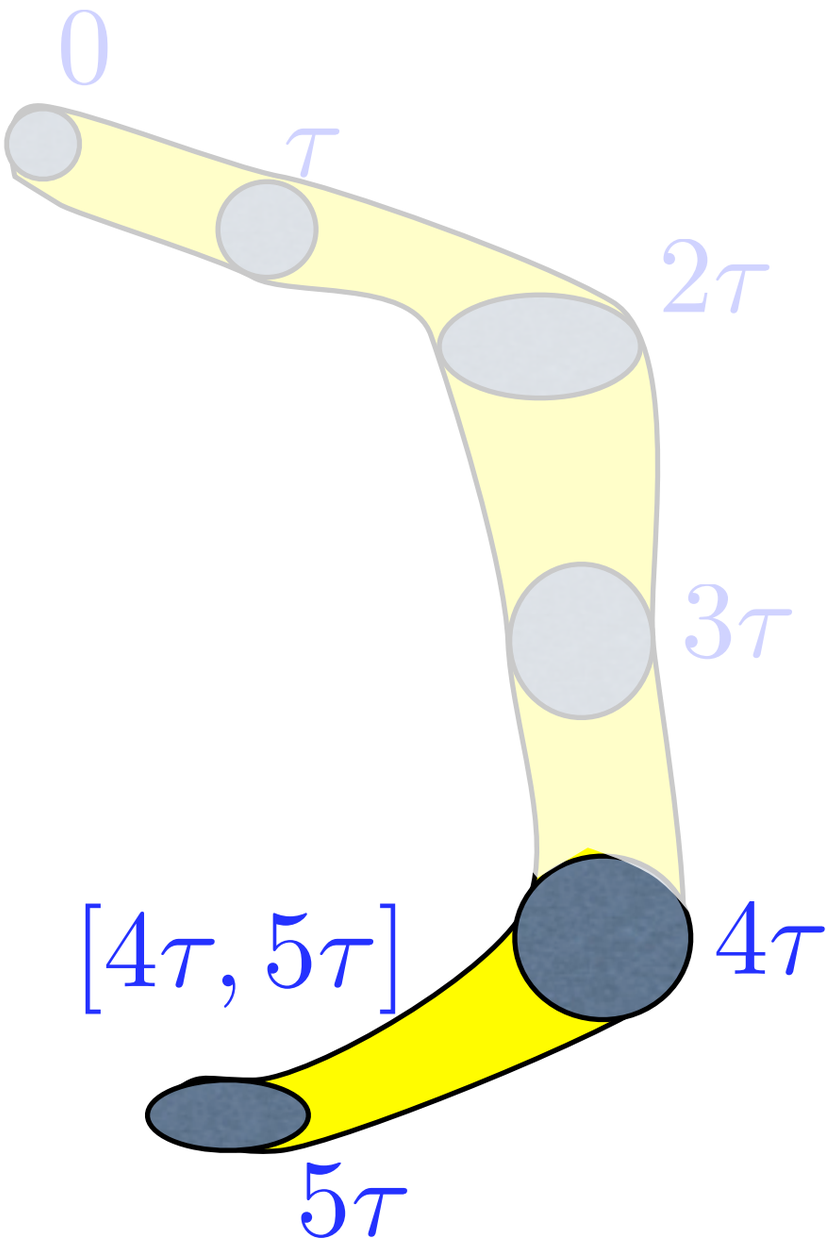


129

Time Interval Discretization for the FSP

★ Instead:

✴ Discretize the time interval into smaller steps and solve a separate projection for each interval.



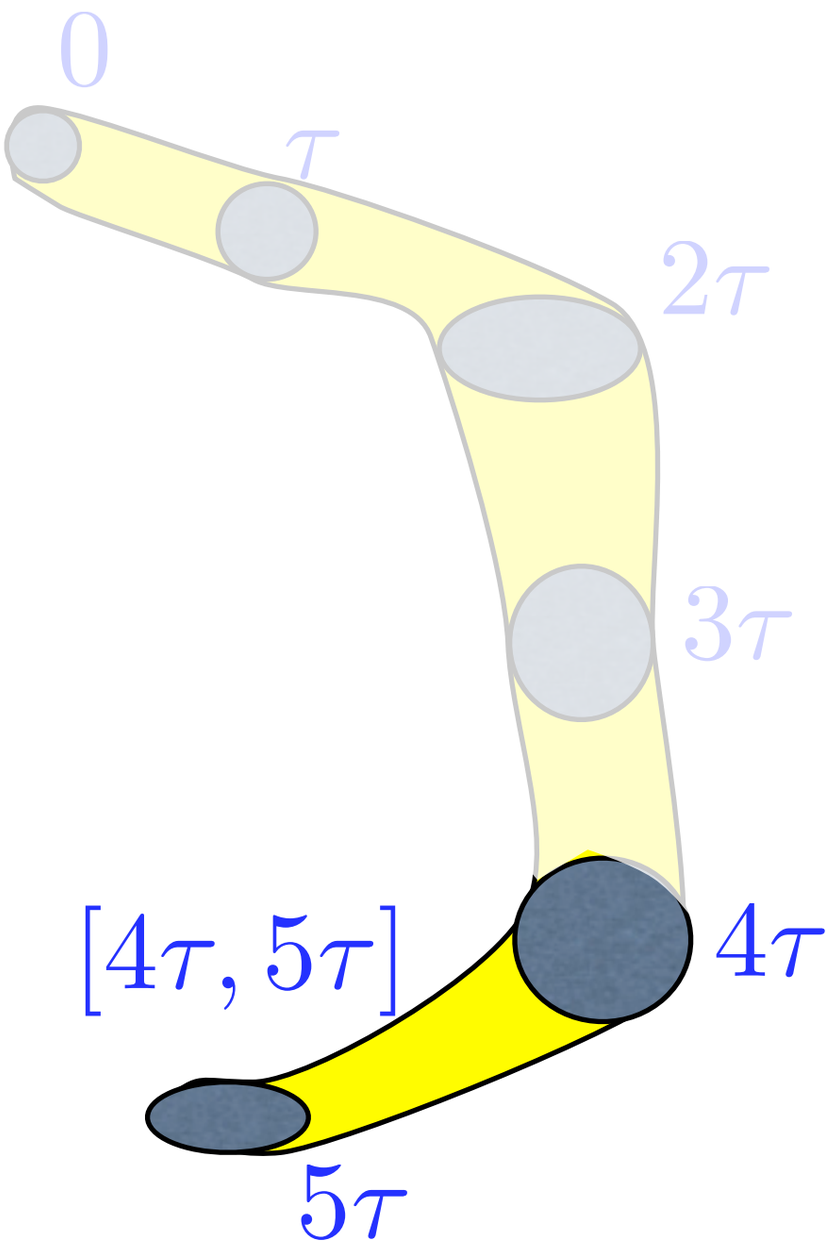
130

Time Interval Discretization for the FSP

★ Solving a few smaller systems can be much easier than solving a single large system.

★ Control the error at each step to obtain a guaranteed final error.

★ Caching and reusing information from one step to the next may further reduce effort.



131

Outline

- ☑ Finite State Projection (FSP)
- Reductions to the FSP
  - ★ Aggregating unobservable states
  - ★ Time interval discretization
  - ★ Slow manifold projection

    Peles/Munsky/Khammash, *JCP*, 2006
  - ★ Coarse meshes for the CME.

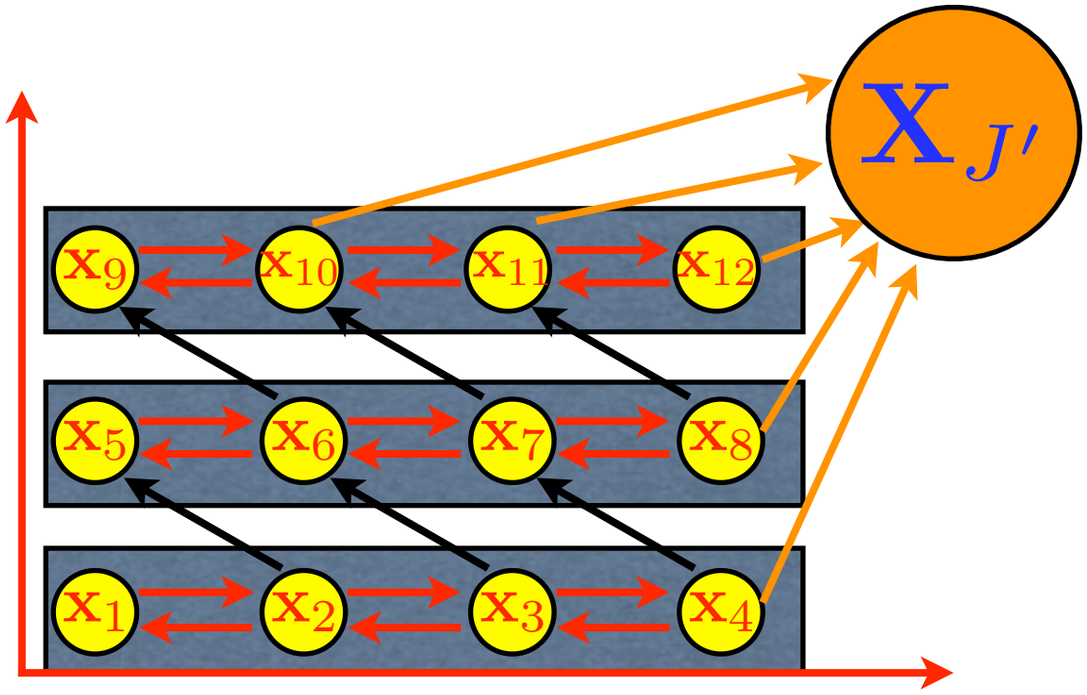
Perturbation Theory and the FSP

- Some reactions occur faster and more frequently than others.
- This can result in a separation of time-scales in the CME.
  - Disadvantages: Often results in numerical stiffness and increased computational complexity.
  - Advantage: May be able to apply perturbation theory to reduce computational effort.

133

Intuition (Slow Manifold Projection)

1. Begin with a finite state (projected) Markov process.
2. Group states connected by frequent reactions.



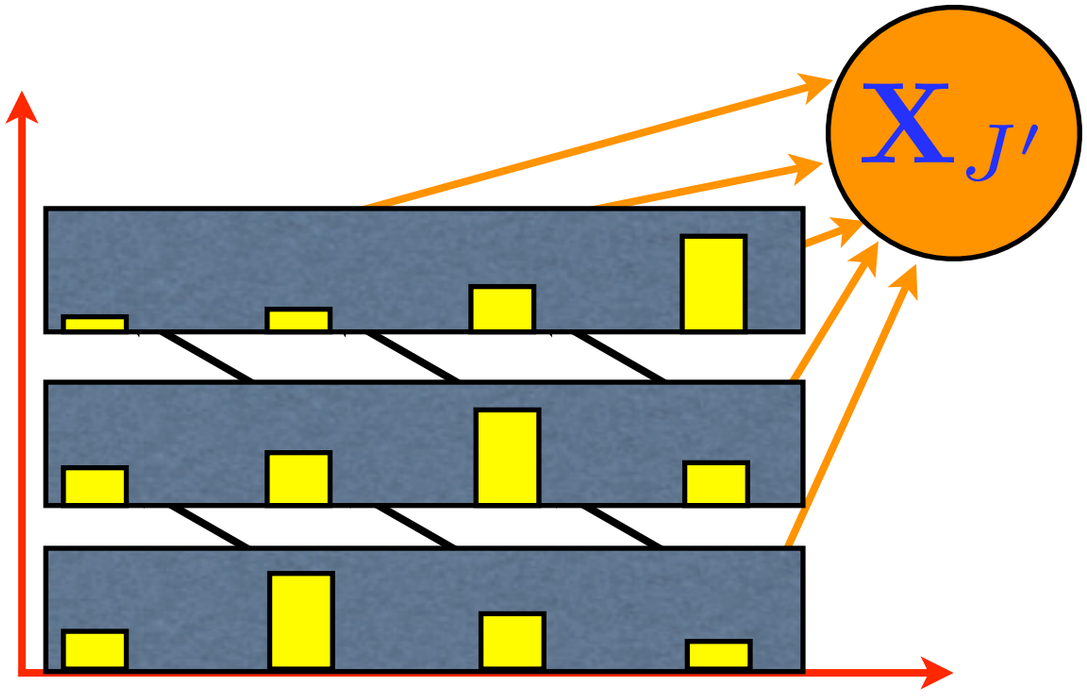
Red Arrows = Fast (Frequent) Reactions

Black Arrows = Slow (Rare) Reactions

Orange Arrows = (Rare) Transitions to Sink

Intuition (Slow Manifold Projection)

1. Begin with a finite state (projected) Markov process.
2. Group states connected by frequent reactions.
3. Find invariant distribution for each group.



Red Arrows = Fast (Frequent) Reactions

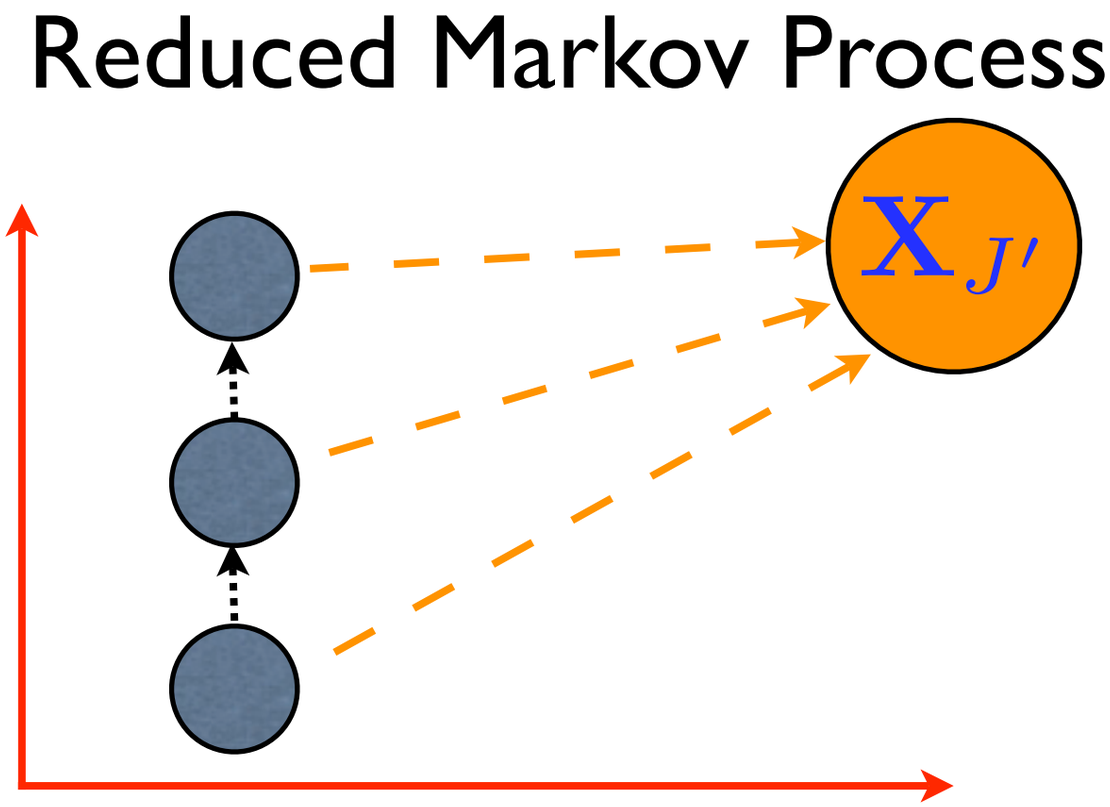
Black Arrows = Slow (Rare) Reactions

Orange Arrows = (Rare) Transitions to Sink

135

Intuition (Slow Manifold Projection)

1. Begin with a finite state (projected) Markov process.
2. Group states connected by frequent reactions.
3. Find invariant distribution for each group.
4. Average to find the rates of the slow reactions.



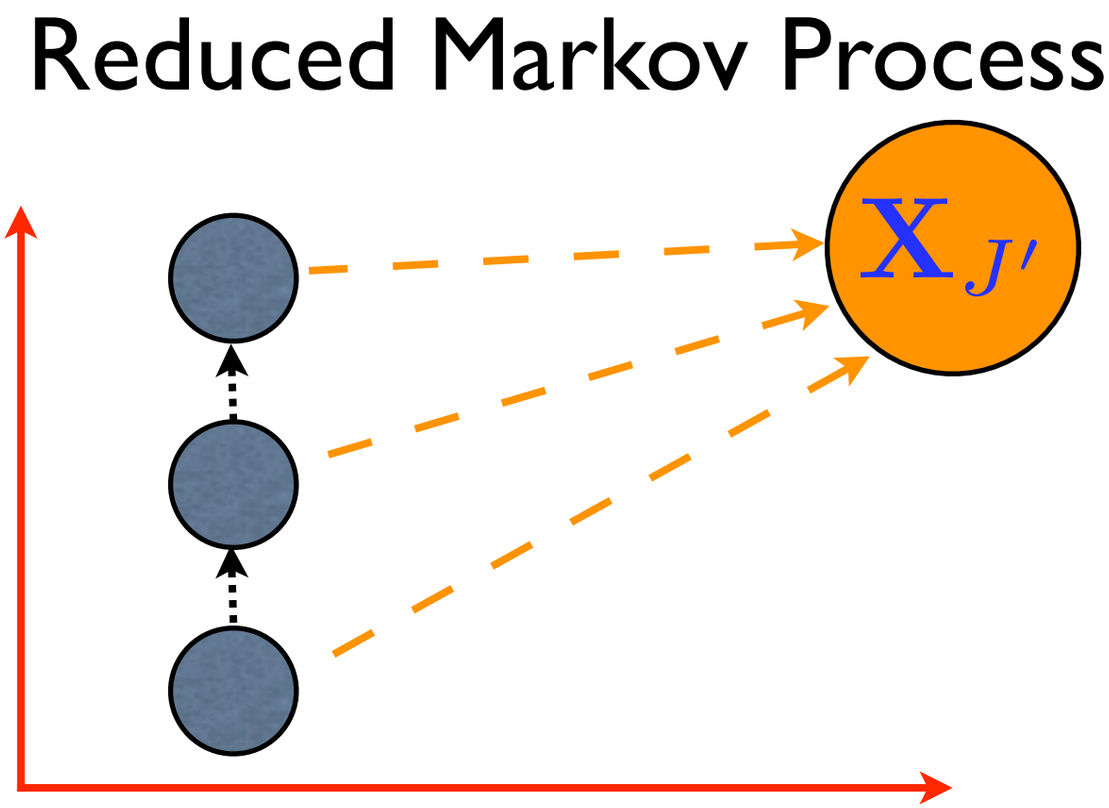
Dotted Black = Averaged Slow Reactions

Dashed Orange = Averaged Transitions to Sink

136

Intuition (Slow Manifold Projection)

1. Begin with a finite state (projected) Markov process.
2. Group states connected by frequent reactions.
3. Find invariant distribution for each group.
4. Average to find the rates of the slow reactions.
5. Solve for the solution on the slow-manifold.
6. Lift solution to original coordinate system.



137

Outline

- ☑ Finite State Projection (FSP)
- Reductions to the FSP

★ Aggregating unobservable states
★ Time interval discretization
★ Slow manifold projection
★ Coarse meshes for the CME

Munsky/Khammash, *IEEE Trans. on Auto. Conrol*, 2008

138

Coarse mesh approximation of the CME

- Precision requirements may change for different regions of the configurations space.

  ★ Small populations require great precision.

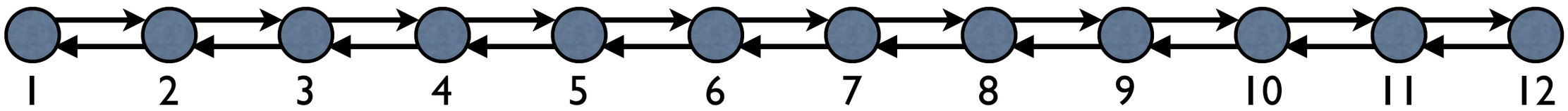
  ★ High populations require far less precision.

- By choosing a good coarse approximation of the CME, we can take advantage of this.

139

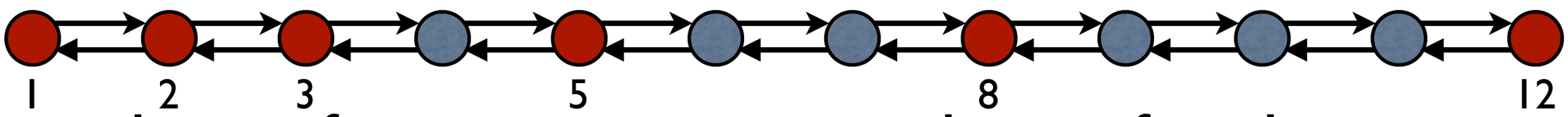
Coarse mesh approximation of the CME

Start with the full 1-dimensional Markov lattice.



$$\dot{\mathbf{P}} = \mathbf{A} \cdot \mathbf{P}(t)$$ Original CME

Choose a subset of mesh points.



and specify an approximate relation for the probability of the removed points: $\mathbf{P} \approx \mathbf{\Phi q}(t)$

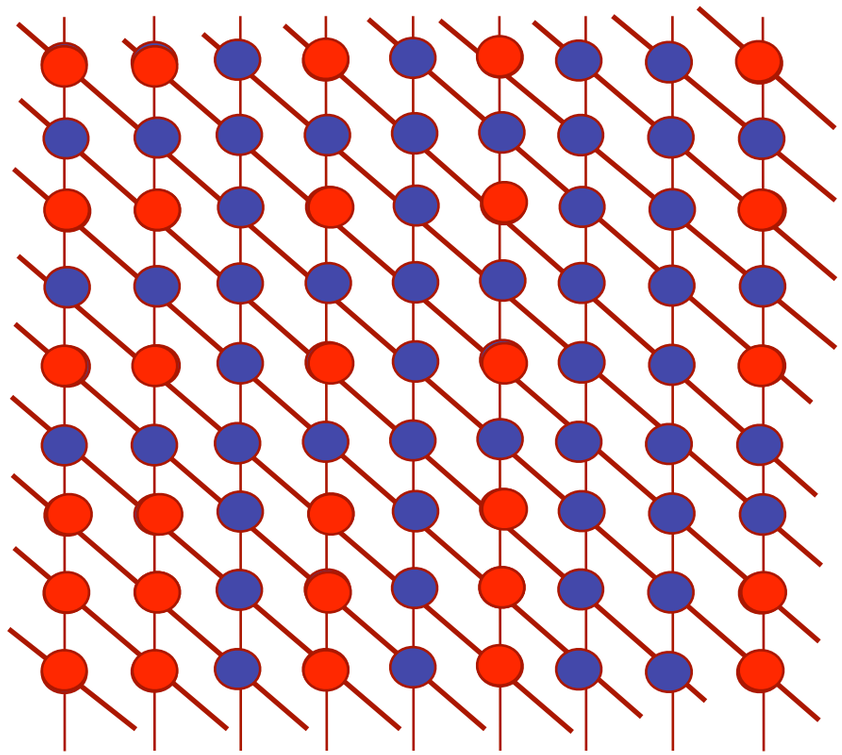
Solve the reduced system ODE: $\dot{\mathbf{q}} = \mathbf{\Phi}^{-L}\mathbf{A}\mathbf{\Phi q}(t)$

and lift back to the original system coordinates:

$$\mathbf{P}(t) \approx \mathbf{\Phi}\exp(\mathbf{\Phi}^{-L}\mathbf{A}\mathbf{\Phi}t)\mathbf{\Phi}^{-L}\mathbf{P}(0)$$

Coarse Mesh:
Multiple-species problems.

1. Begin with original lattice.
2. Choose interpolation points.
3. Form interpolation (shape) function: $\mathbf{P}(t) \approx \mathbf{\Phi q}(t)$
4. Project system to find reduced system of ODEs:
   $\dot{\mathbf{q}}(t) = \mathbf{\Phi}^{-L}\mathbf{A\Phi q}(t)$
5. Solve reduced system.
6. Lift back to original coordinates.



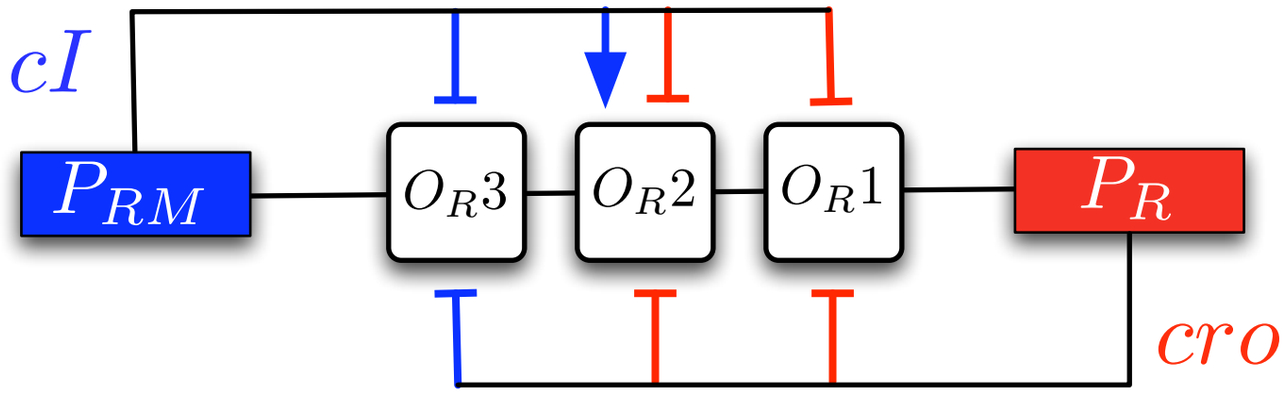
141

Outline

- ☑ Finite State Projection (FSP)
- ☑ Reductions to the FSP
- Case Studies
  - ★ Lambda Phage.
  - ★ Heat Shock.

142

A toy model of phage lambda

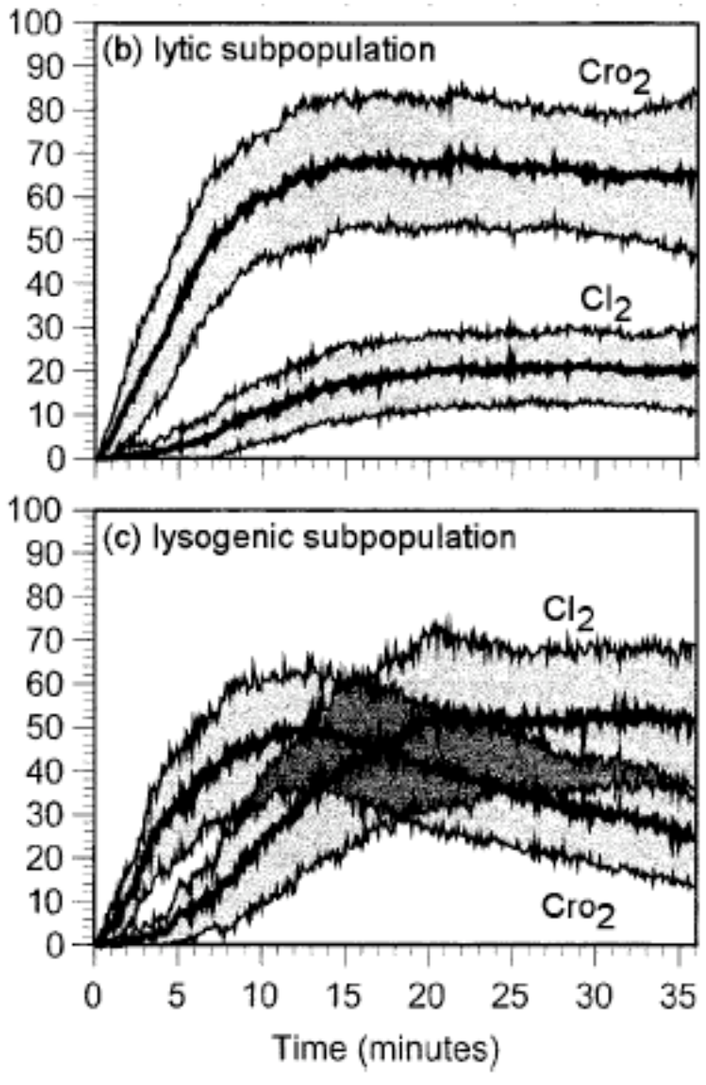


- We consider only the core of the lambda switch.
- Two proteins, cI and cro.
- These activate and repress the P_R and P_{RM} promoters according to the model of Shea and Ackers, 1985.

143

The Phage Lambda Lysis-Lysogeny Decision

Arkin, Ross, McAdams, 1998.
Full Model



Lytic fate

★ Cro reaches a high level before CI is produced in much quantity.

★ Cro represses transcription of CI.

Lysogenic fate

★ CI increases a little earlier.

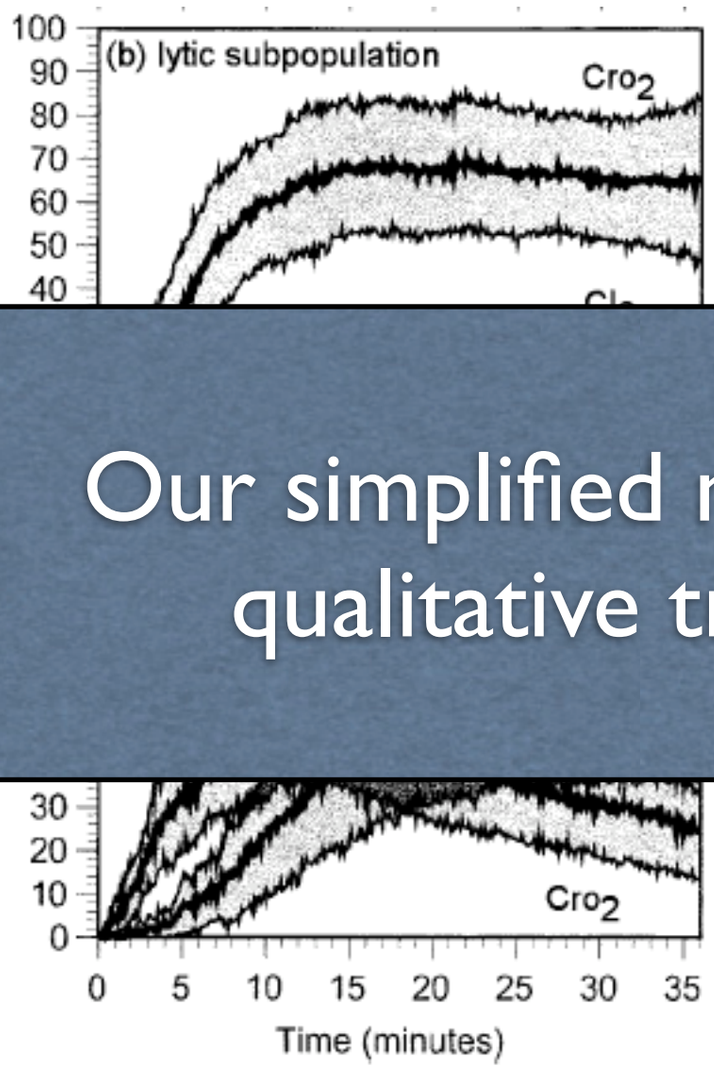
★ CI represses transcription of Cro.

★ CI is free to increase even further.

144

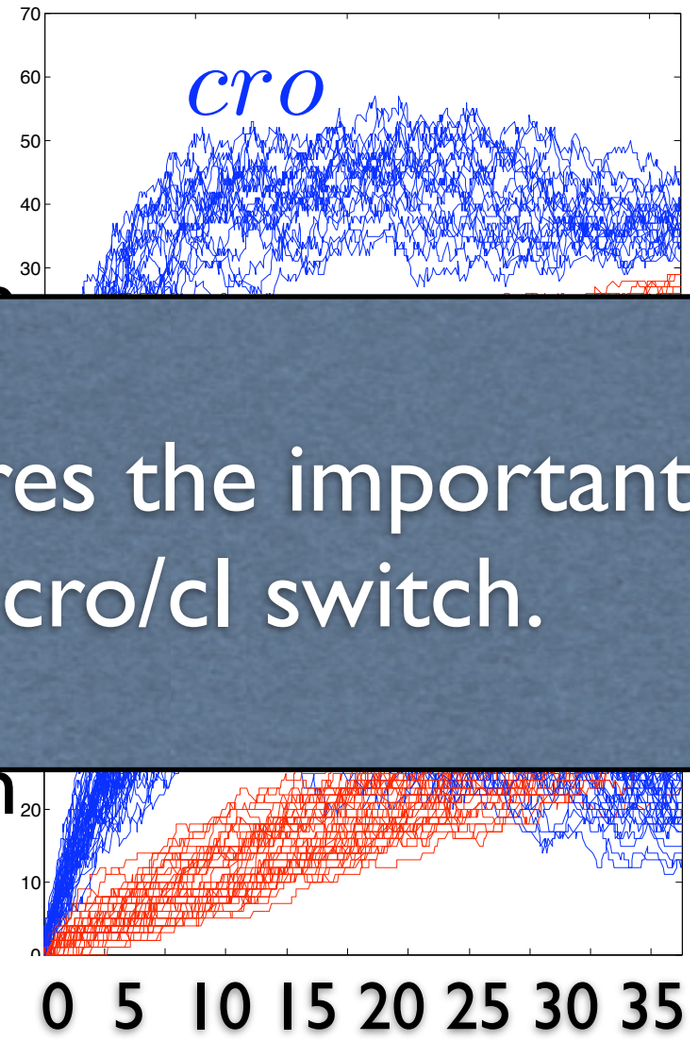
Relevance of Current Model

Arkin, Ross, McAdams, 1998.
Full Model



Lytic subpopulation

Current simplified model

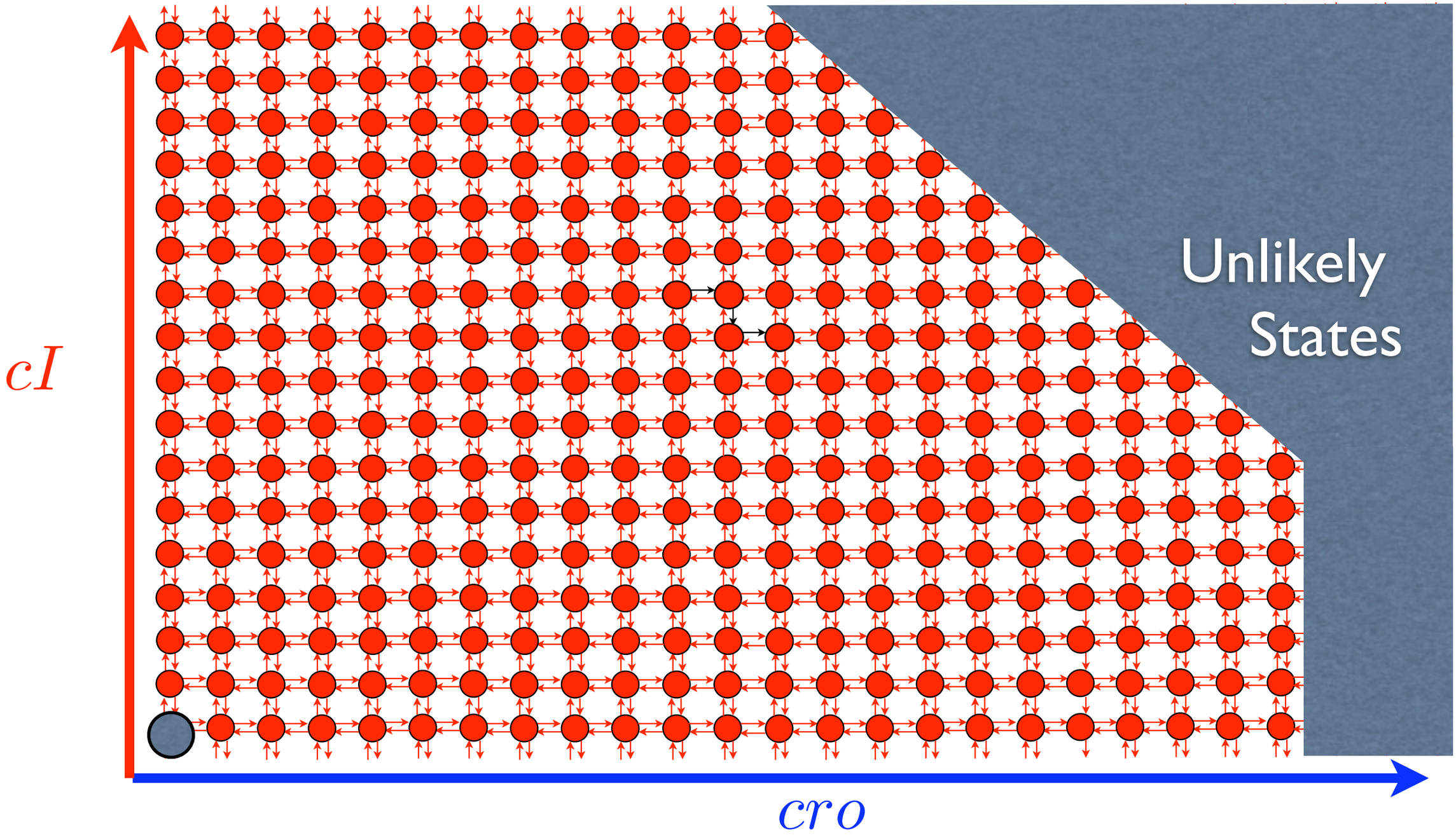


Our simplified model captures the important qualitative trends of the cro/cI switch.

Computations done using Gillespie's SSA.

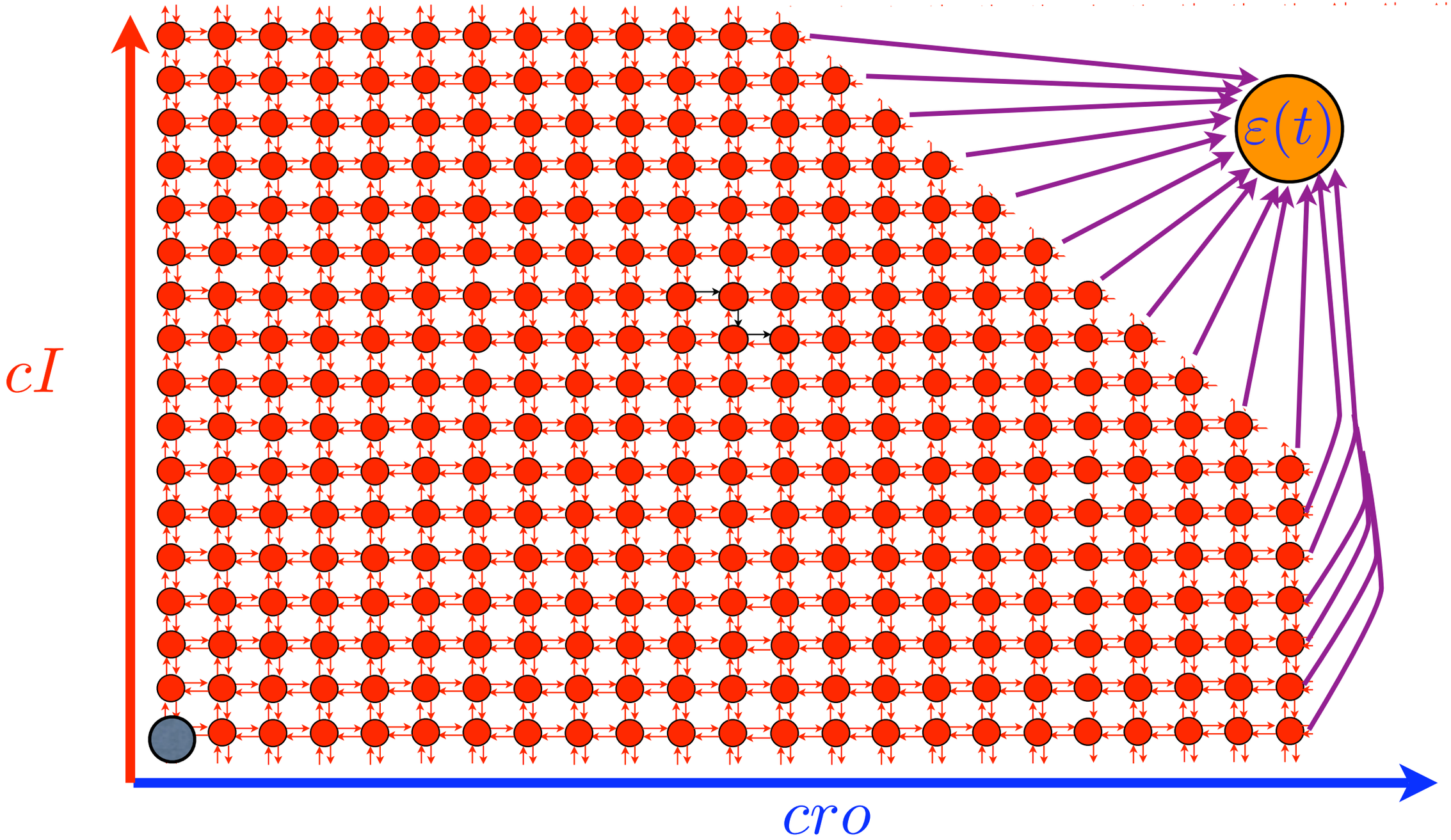
145

Applying the FSP to the Phage Lambda Switch



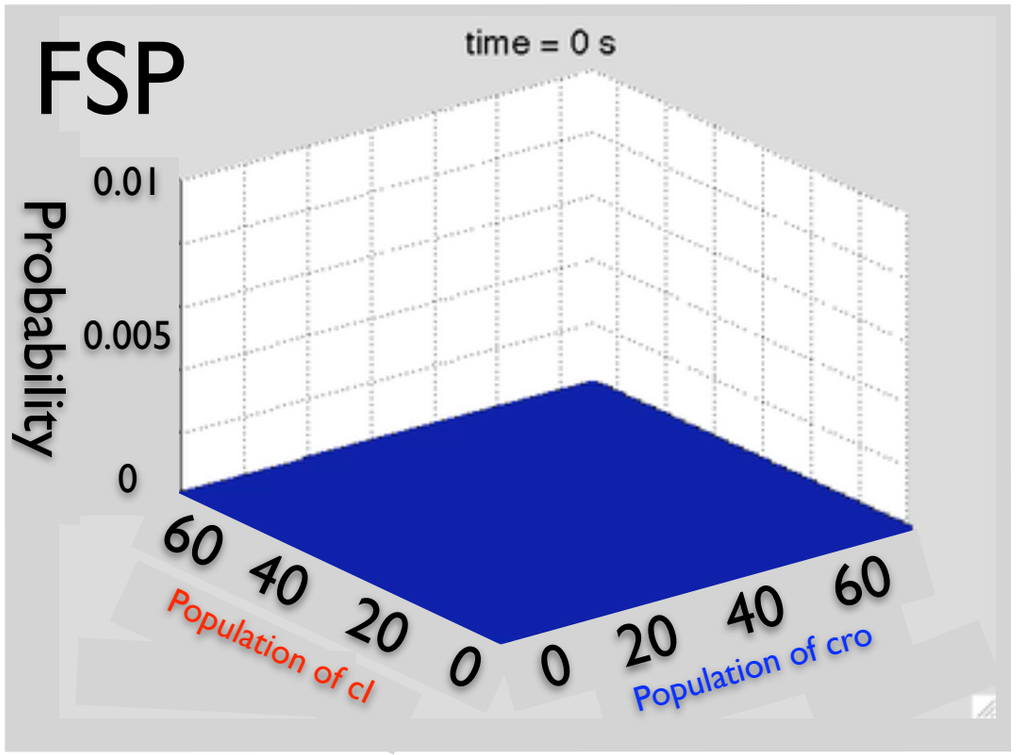
146

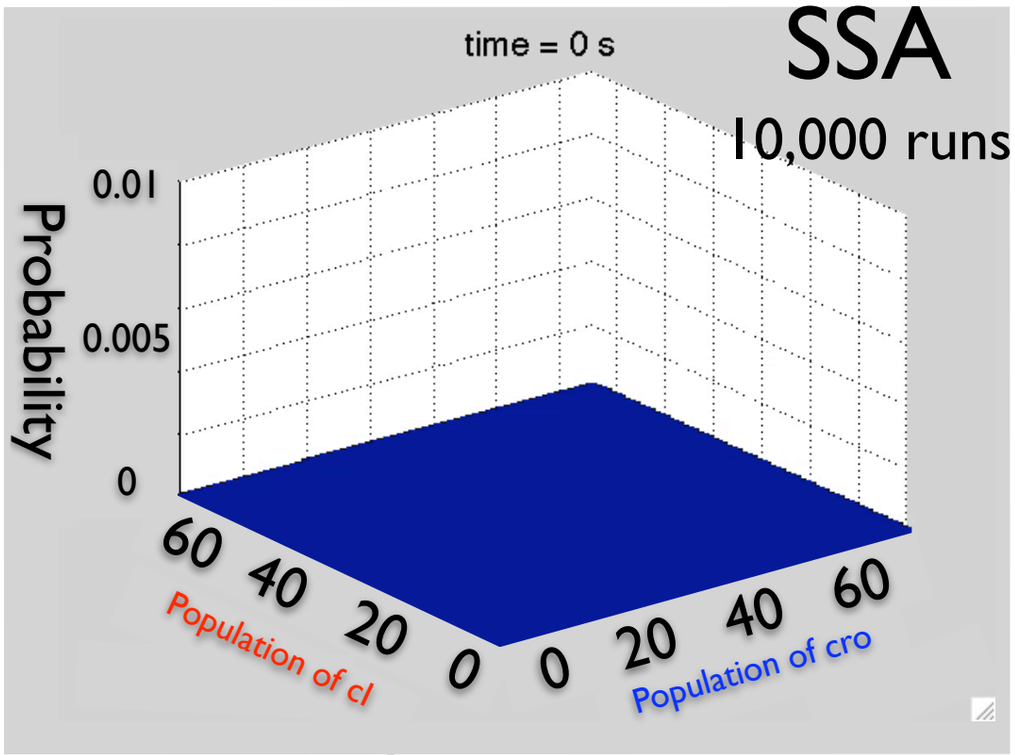
Applying the FSP to the Phage Lambda Switch



147

Efficiency and Accuracy of FSP Results





Method	# Simulations	Time (s)	$\|\|\text{Error}\|\|_1$
FSP	– [a]	163	$\leq 5.3 \times 10^{-3}$

Guaranteed

No Guarantees

[a] The FSP algorithm is run only once.

148

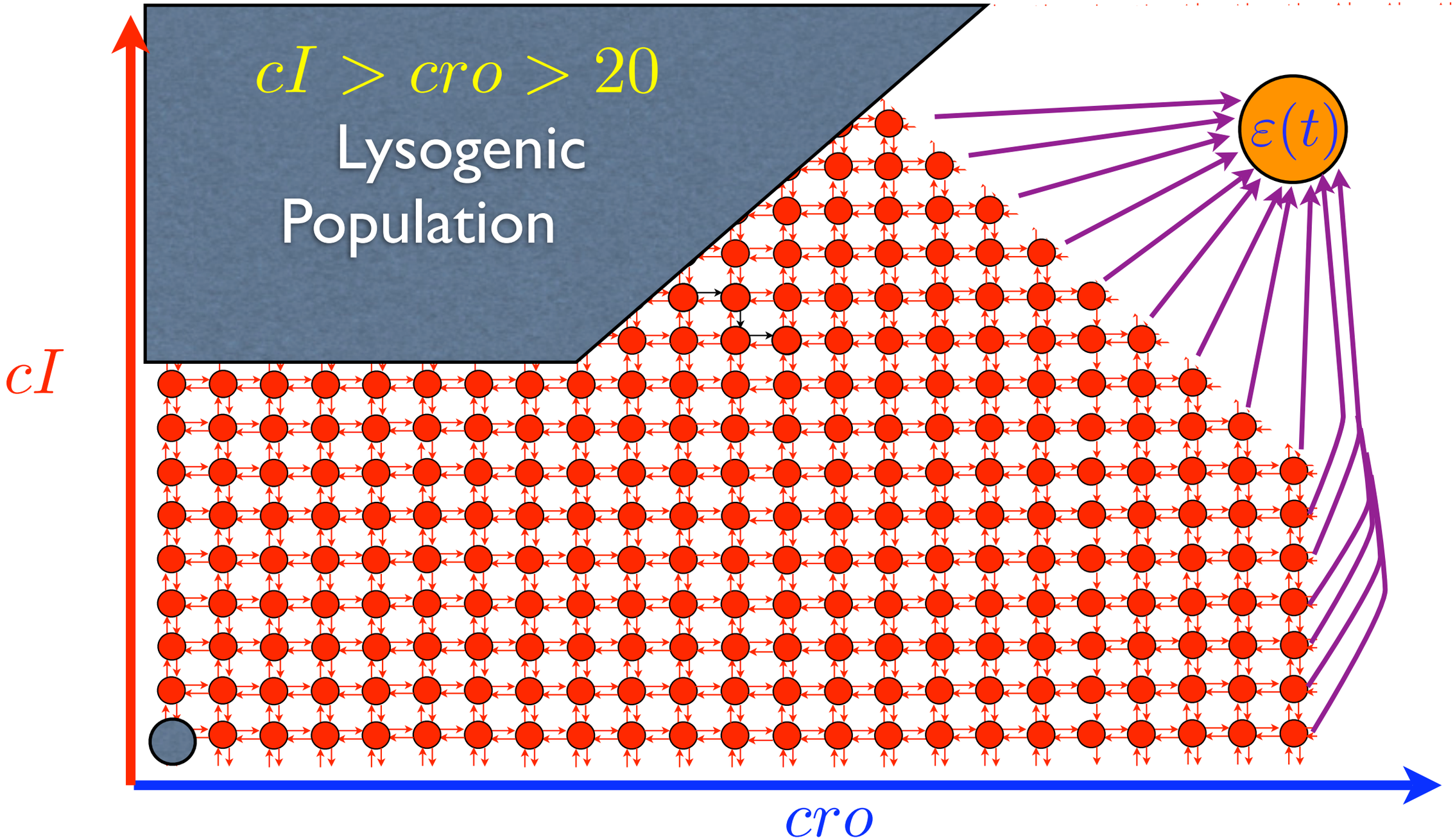
Additional information available with the FSP solution

- In many cases the FSP is faster and more accurate the Monte Carlo methods.
- Higher precision allows greater flexibility.

★ Direct Computation of Switch Rates.

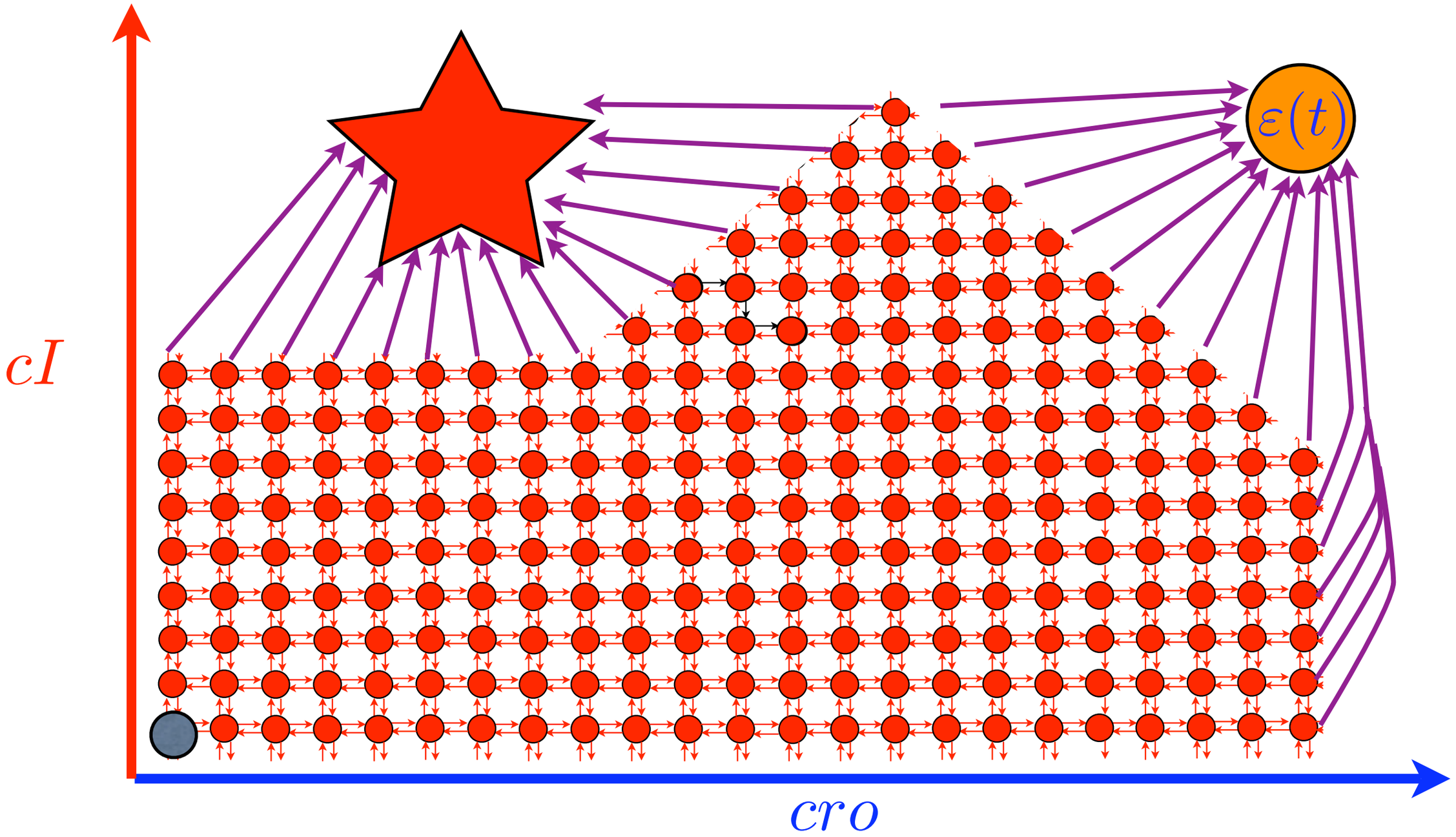
149

Using the FSP to Compute Switch Rates



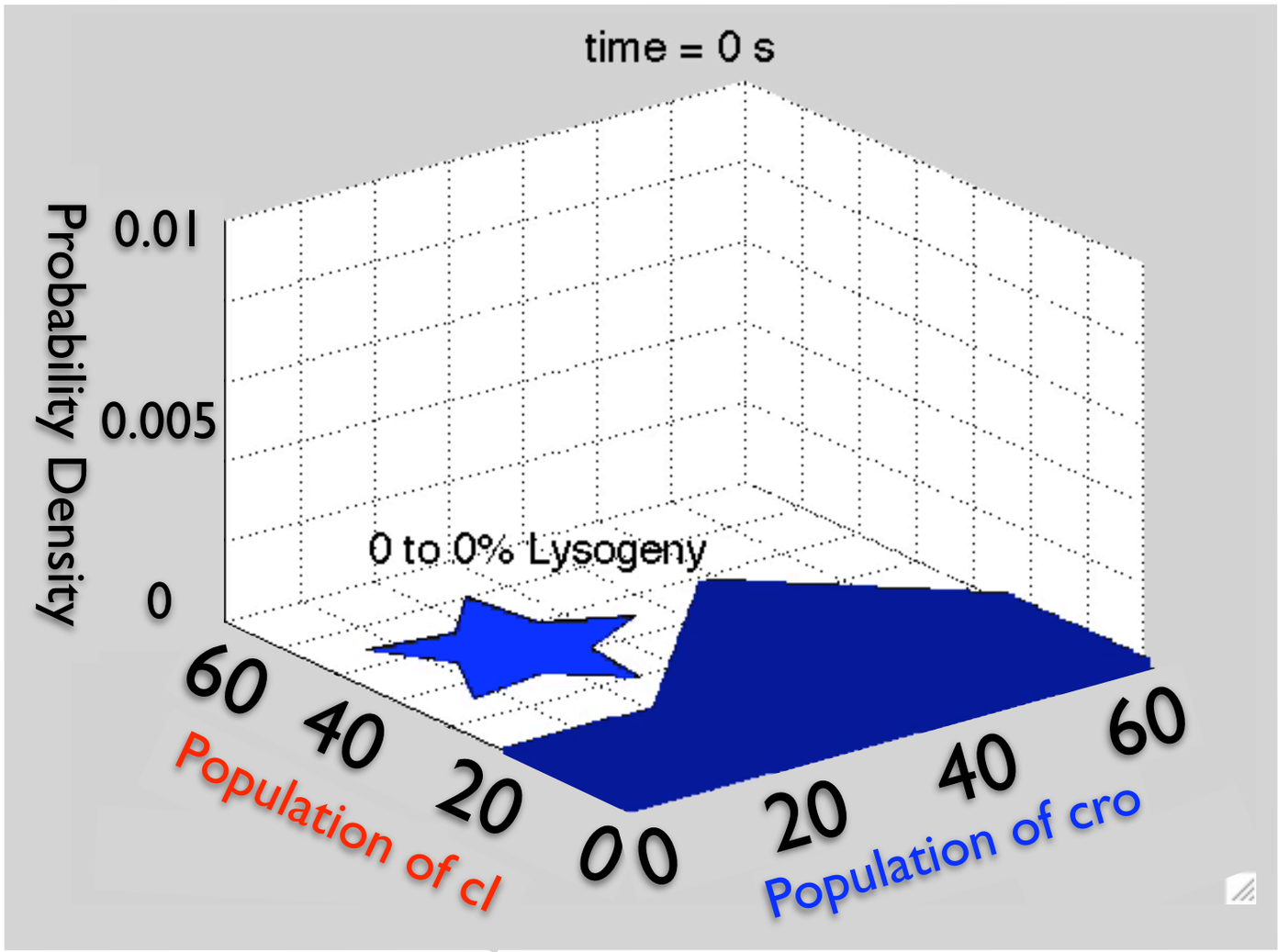
150

Using the FSP to Compute Switch Rates



151

Using the FSP to Compute Switch Rates



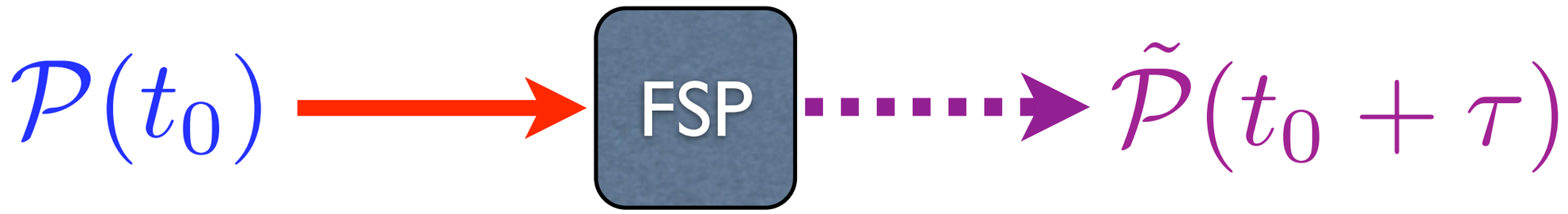
Method	Time (s)	Relative Error	Guarantee?
FSP	25.5 s	< 0.08 %	yes
10^4 SSA runs	440.0 s	≈ 0.90 %	no

Additional information available with the FSP solution

- In many cases the FSP is faster and more accurate the Monte Carlo methods.
- Higher precision allows greater flexibility.
  - ★ Direct Computation of Switch Rates.
  - ★ Simultaneous consideration of many different initial conditions.

153

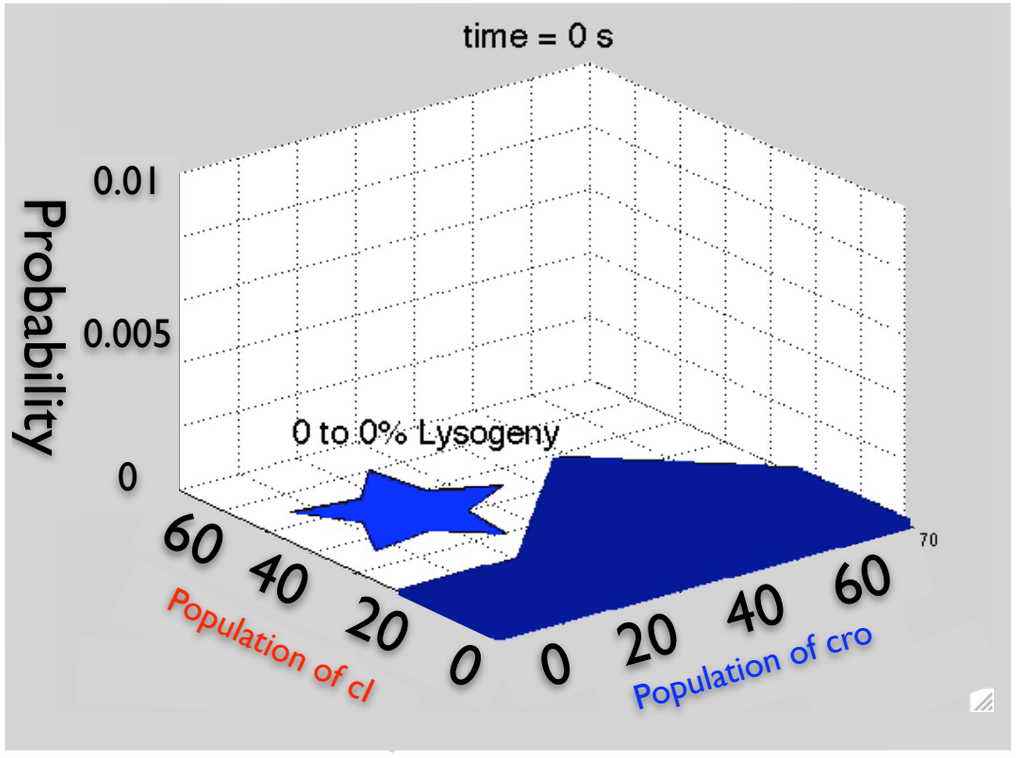
Comparing different initial conditions.



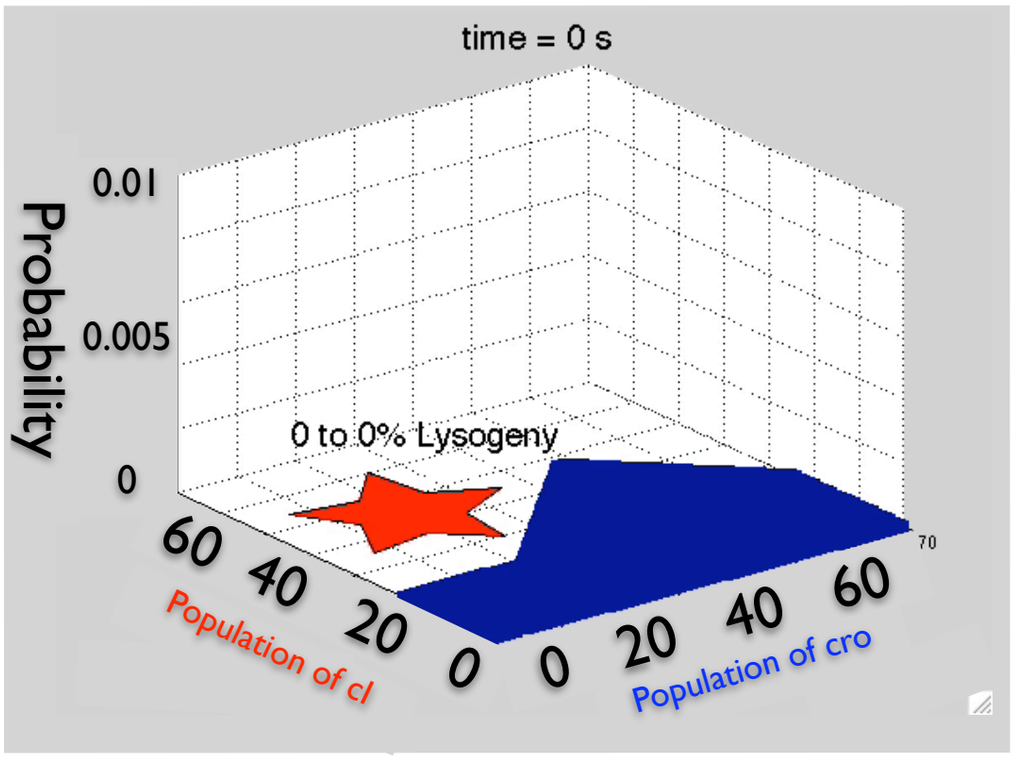
- The FSP is an approximate map of distributions from one time to another.
- This map is valid for any initial distribution.
  - ★ Once computed, this map is cheap to apply again and again.
  - ★ The map automatically provides error bounds for any initial condition!

Comparing different initial conditions. (Increase in cro)

$$cI_0 = 0$$
$$cro_0 = 0$$



$$cI_0 = 0$$
$$cro_0 = 5$$

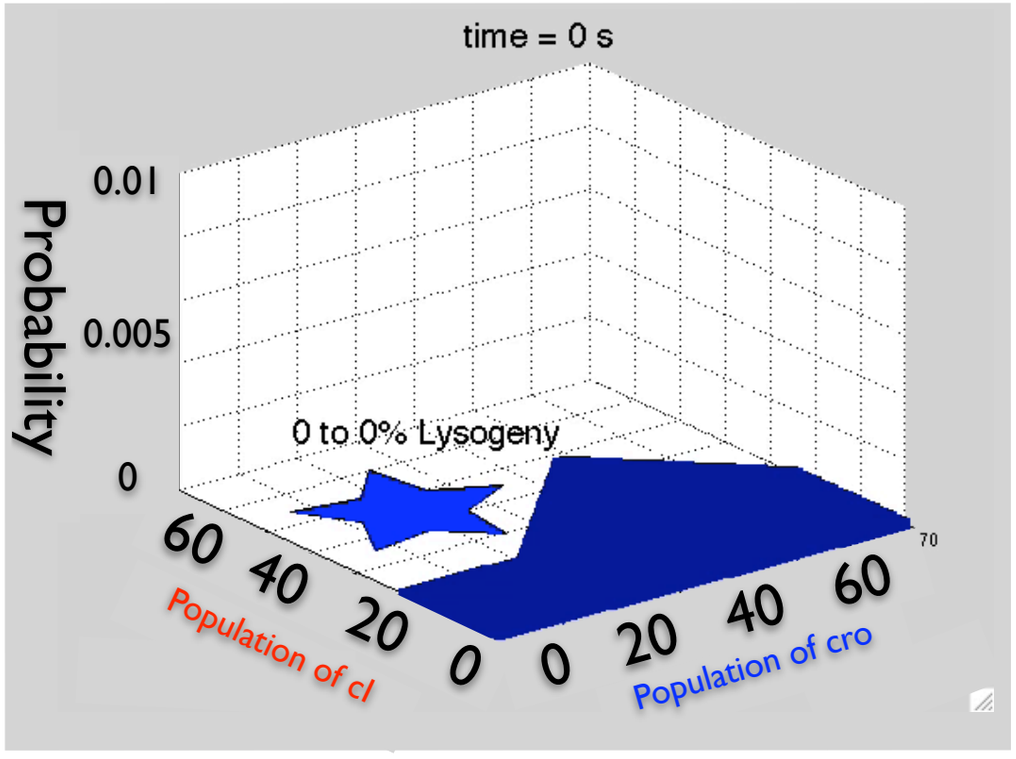


Increasing the initial amount of cro yields a slight decrease in the lysogeny rate.

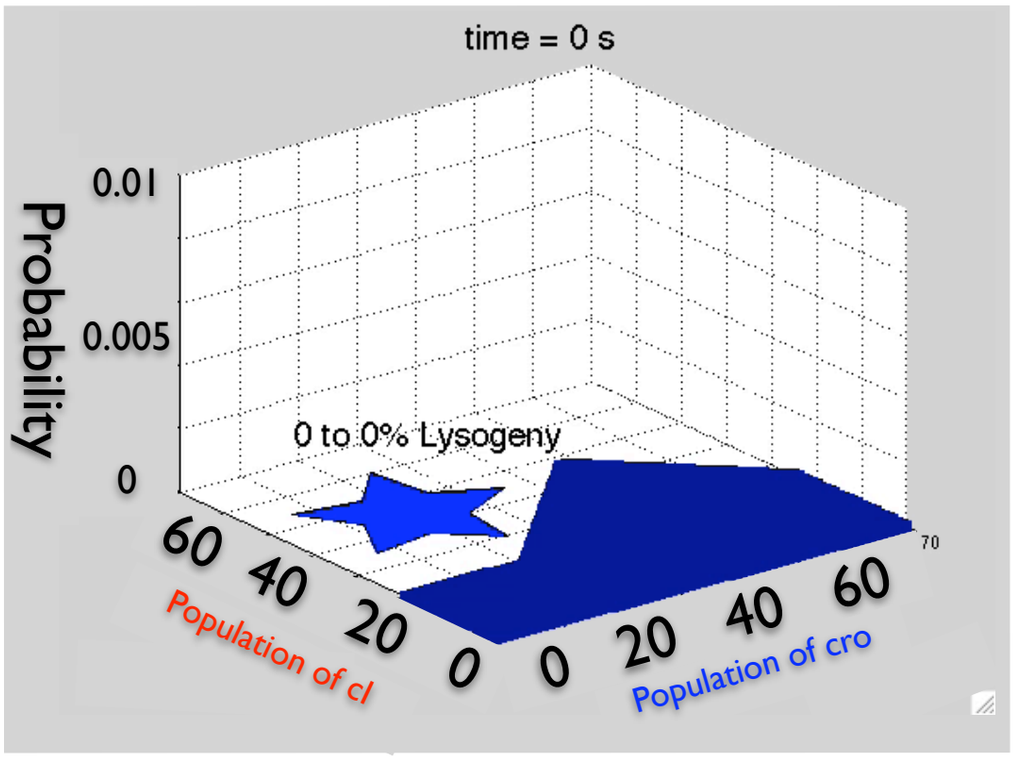
155

Comparing different initial conditions. (Increase in cI)

$$cI_0 = 0$$
$$cro_0 = 0$$



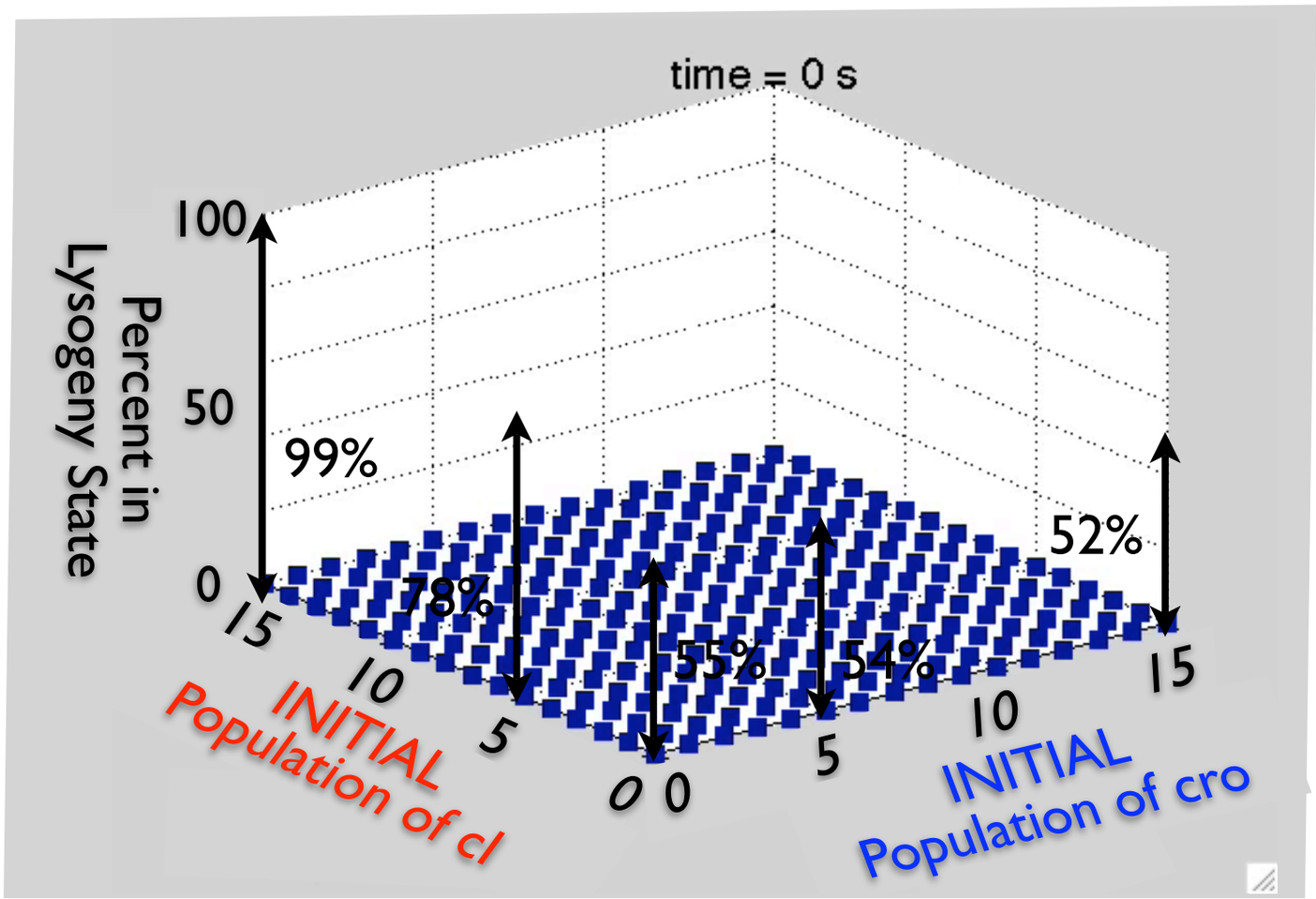
$$cI_0 = 5$$
$$cro_0 = 0$$



Increasing the initial amount of cI yields a significant increase in lysogeny rate.

156

Simultaneous comparison of an array of initial condition.)



Method	Time (s)	# I.C.'s	$\|\|Error\|\|_1$	Guarantee?
FSP	66.9 s	2000	$< 1 \times 10^{-4}$	yes
10^4 SSA runs	440.0 s	1	≈ 0.09	no
10^{13} SSA runs	$\approx$ 14,000 years!	2000	$\approx 1 \times 10^{-4}$	no

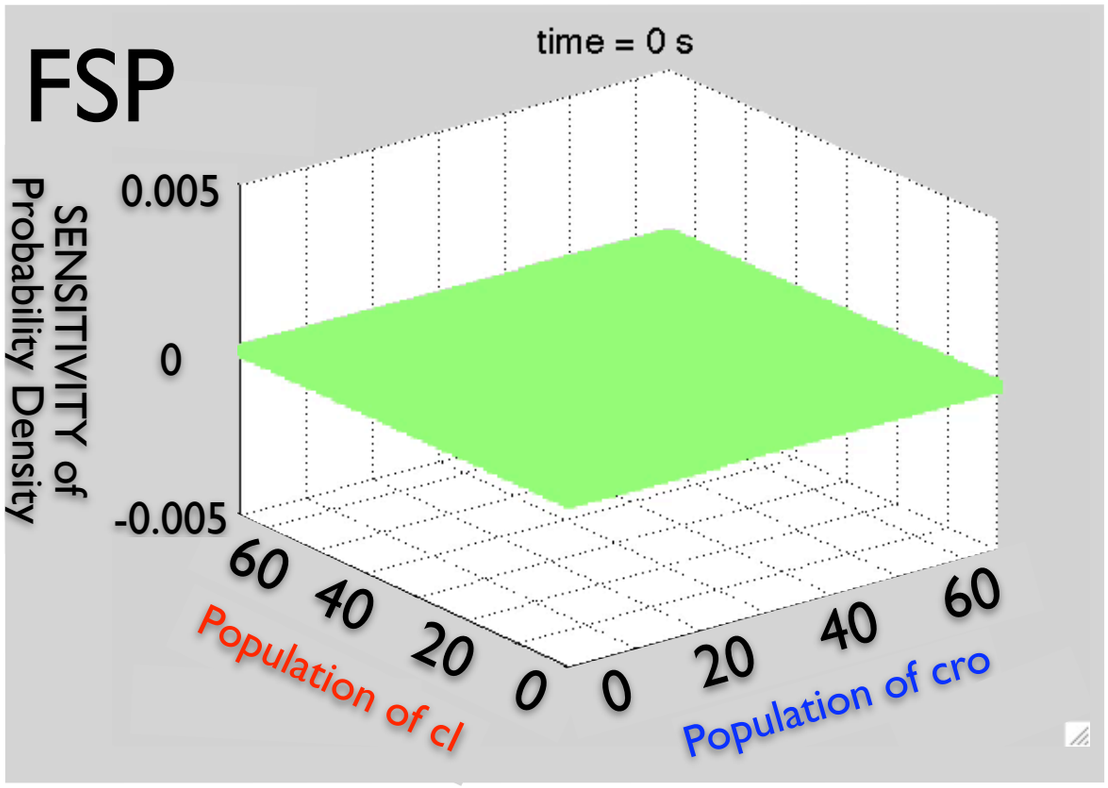
Additional information available with the FSP solution

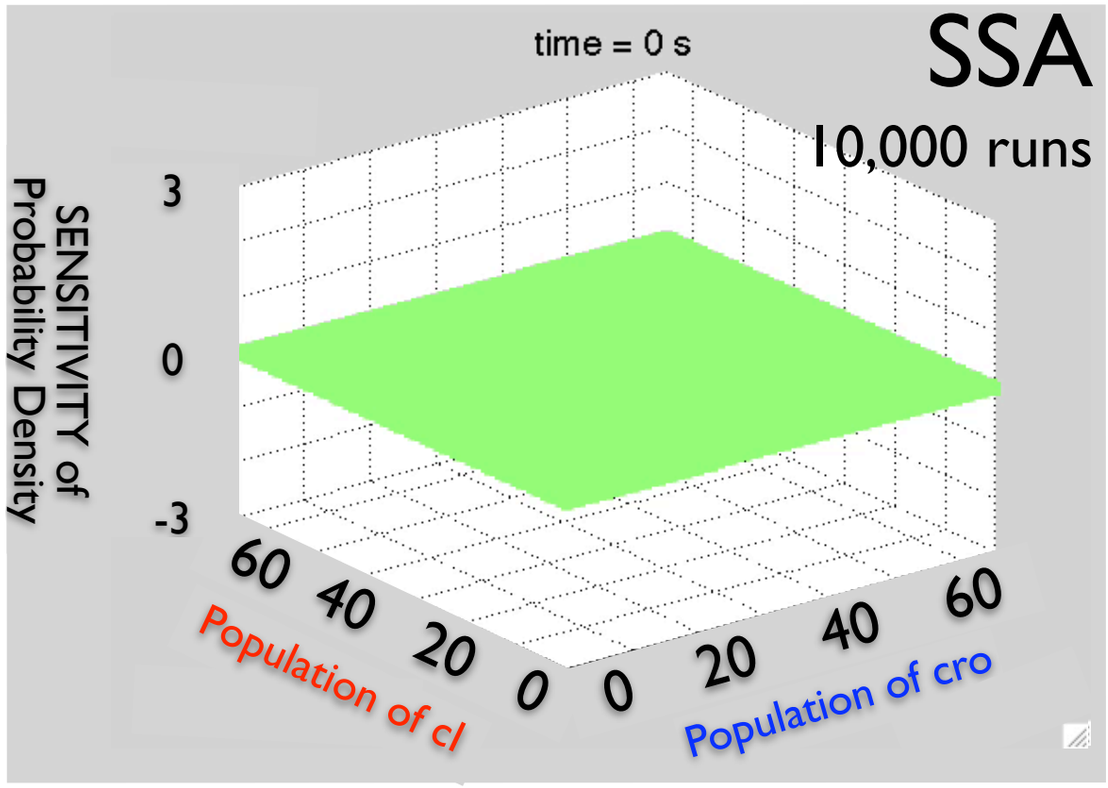
- In many cases the FSP is both faster and more accurate than other available methods.
- Higher precision allows greater flexibility.
  - ★ Direct Computation of Switch Rates.
  - ★ Simultaneous consideration of many different initial conditions.
  - ★ Sensitivity to parameter changes.

158

Parametric Sensitivity of Probability Distributions.

Sensitivity to a small increase in cell Volume.





★ Sensitivity analysis requires a huge degree of accuracy.

★ Monte Carlo methods would require hundreds of millions of runs!!

159

Outline

- ☑ Finite State Projection (FSP)
- ☑ Reductions to the FSP
- Case Studies
  - ★ Lambda Phage.
  - ★ Heat Shock.

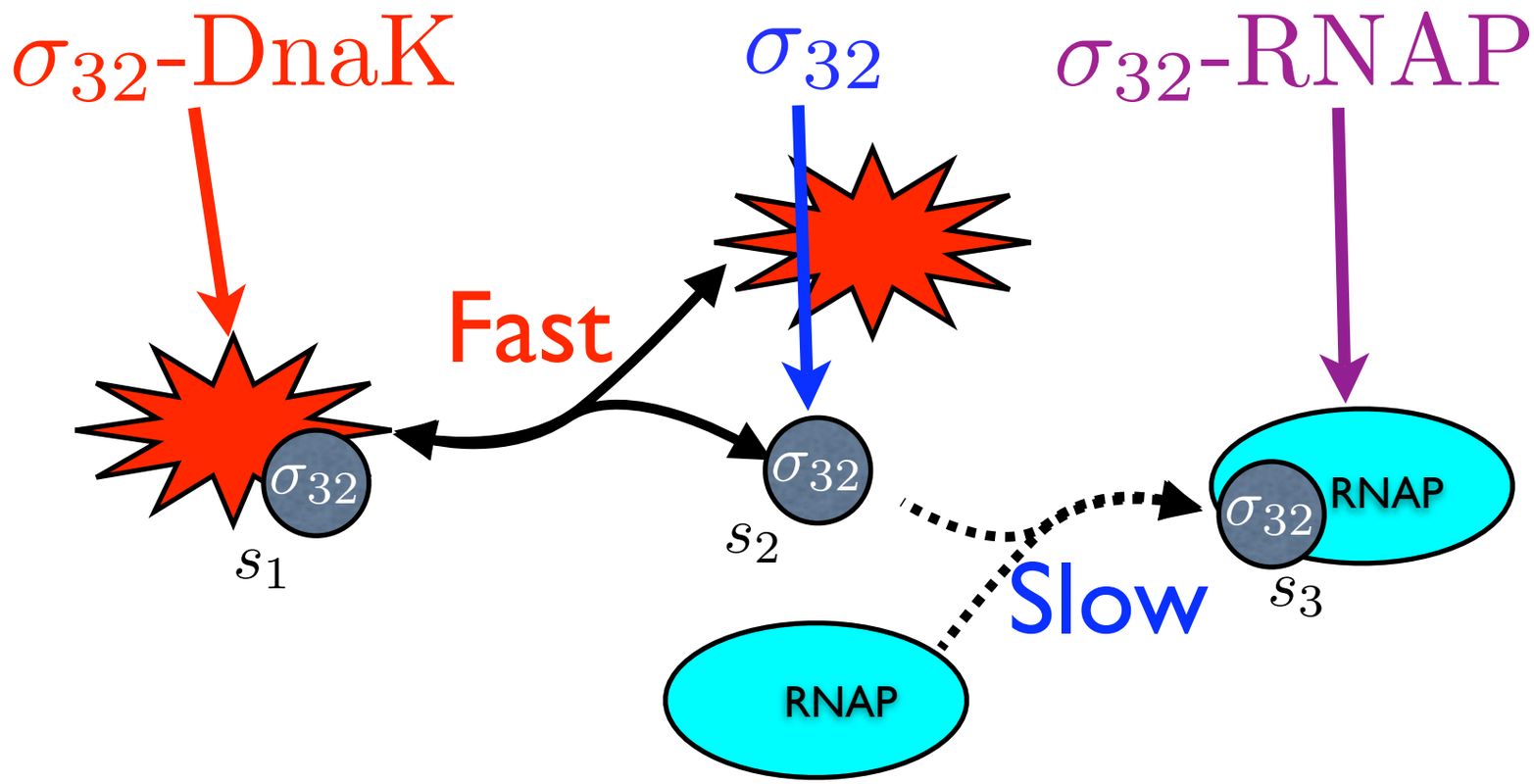
160

Toy Heat Shock Model in *E. coli*

3 forms for σ_{32}:

$$S_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} S_2$$

$$S_2 \xrightarrow{k_3} S_3$$



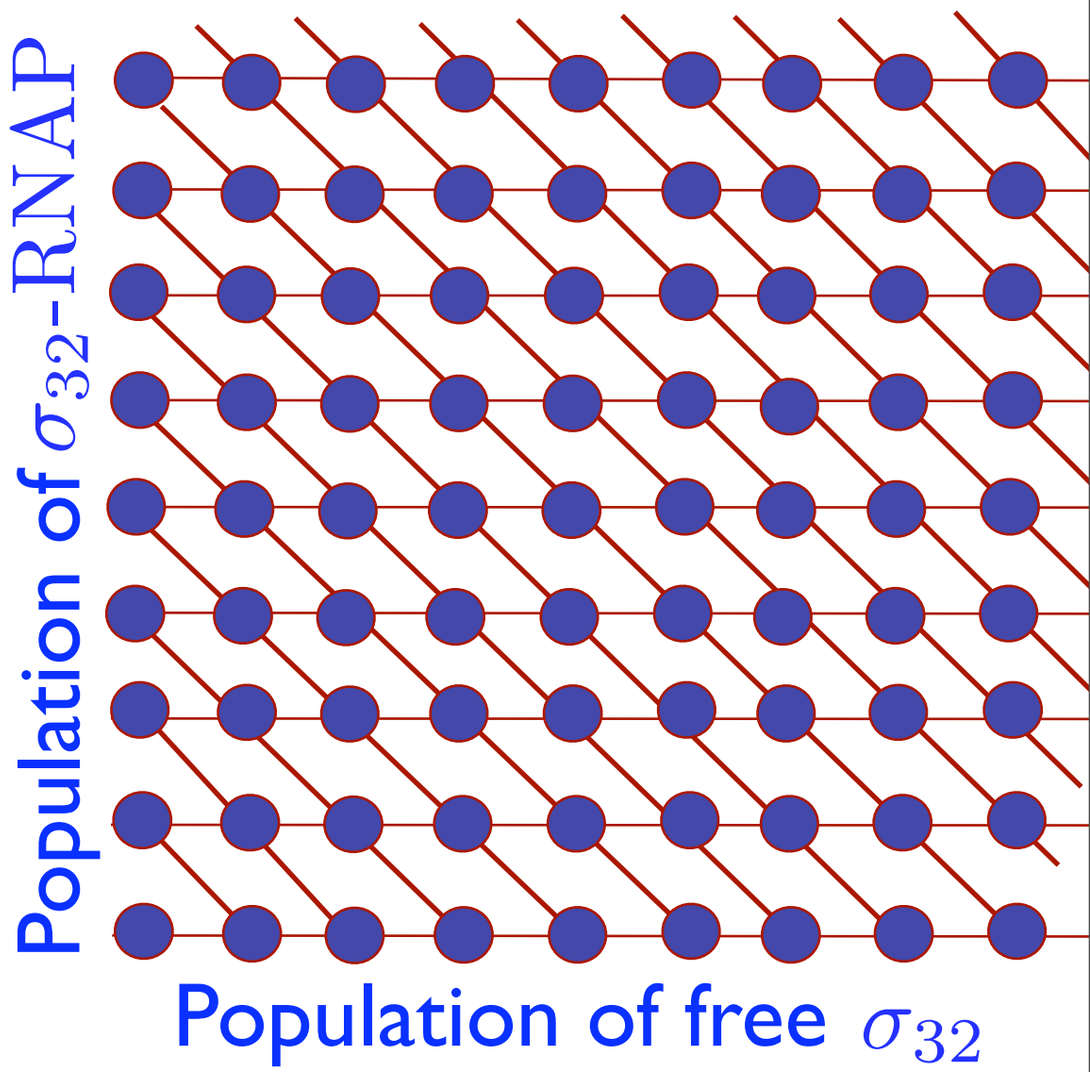
El Samad et al, *PNAS*, vol. 102, No. 8, 2005

161

Toy Heat Shock Model in *E. coli* (cont.)

Five Different FSP Solution Schemes:

1. Full FSP



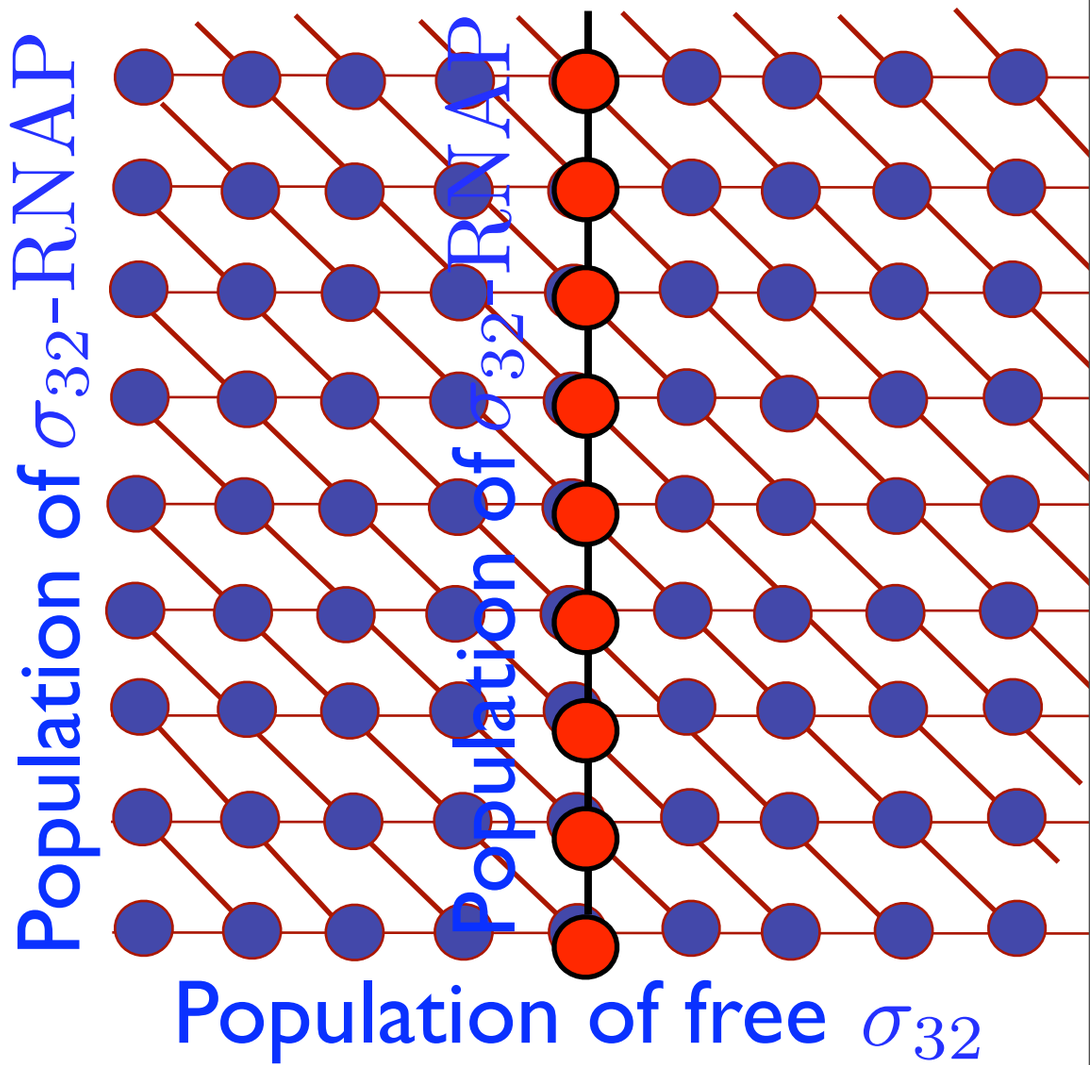
4459 ODEs

162

Toy Heat Shock Model in *E. coli* (cont.)

Five Different FSP Solution Schemes:

1. Full FSP
2. Slow manifold (FSP-SM)



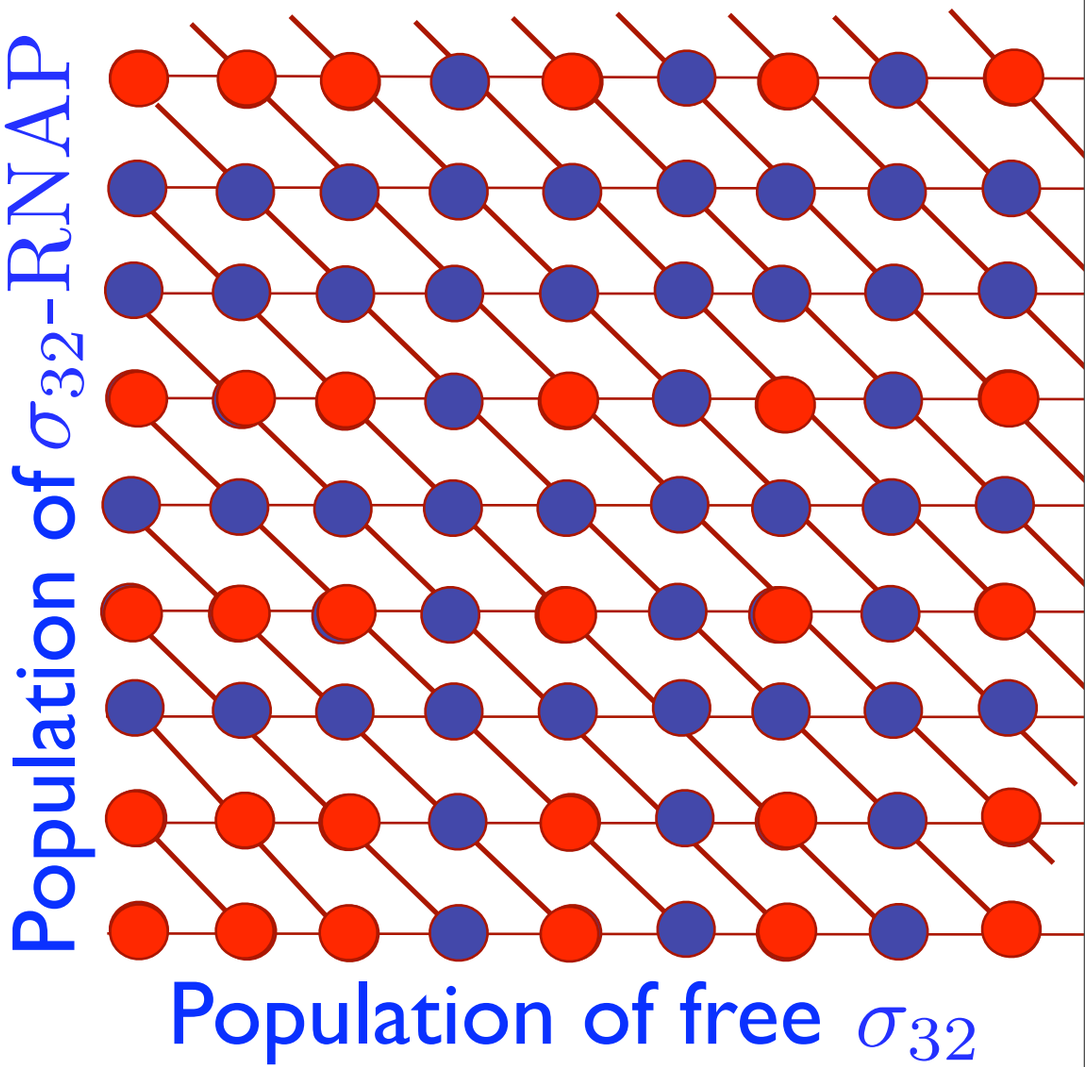
343 ODEs

163

Toy Heat Shock Model in *E. coli* (cont.)

Five Different FSP Solution Schemes:

1. Full FSP
2. Slow manifold (FSP-SM)
3. Interpolated (FSP-I)



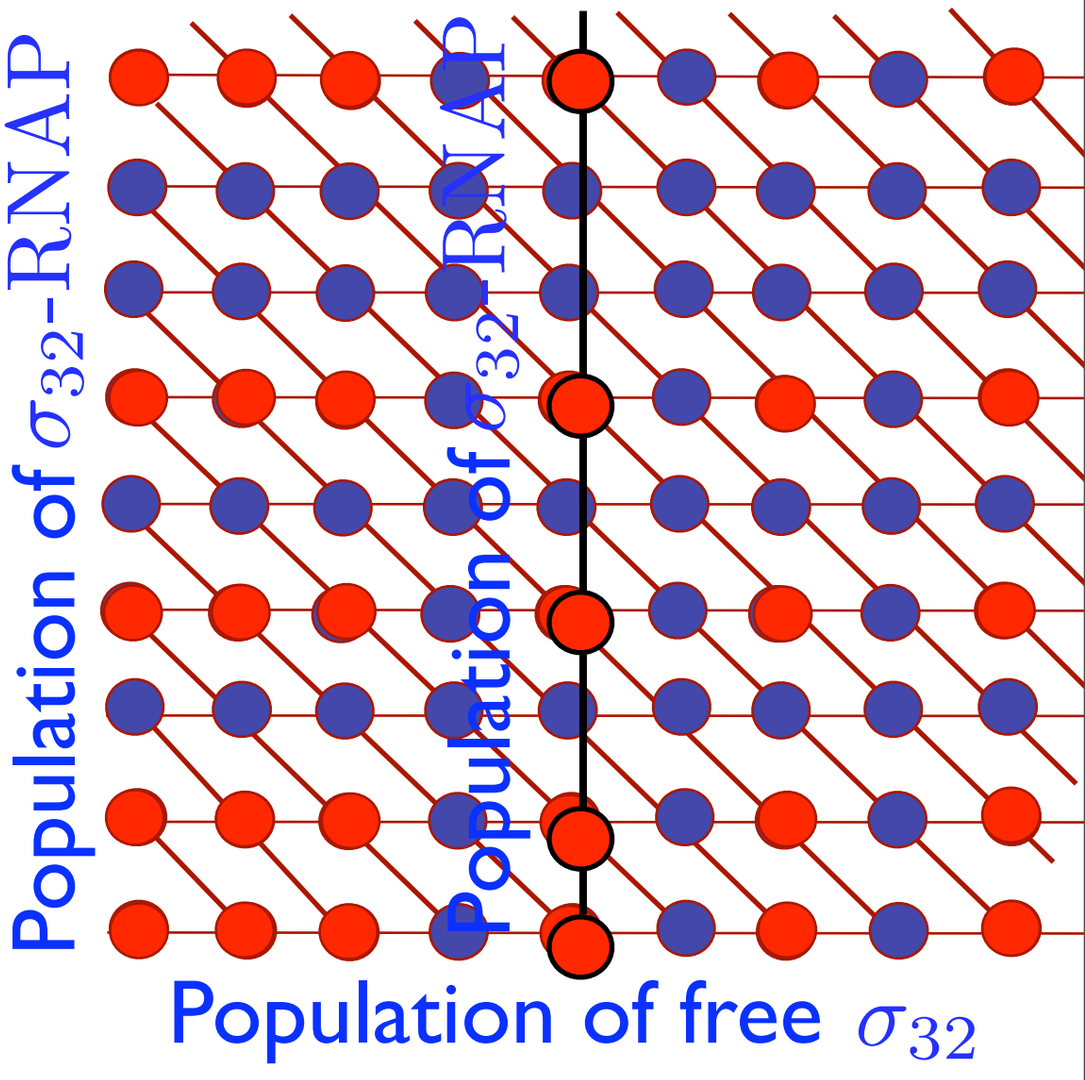
539 ODEs

164

Toy Heat Shock Model in *E. coli* (cont.)

Five Different FSP Solution Schemes:

1. Full FSP
2. Slow manifold (FSP-SM)
3. Interpolated (FSP-I)
4. Hybrid (FSP-SM/I)



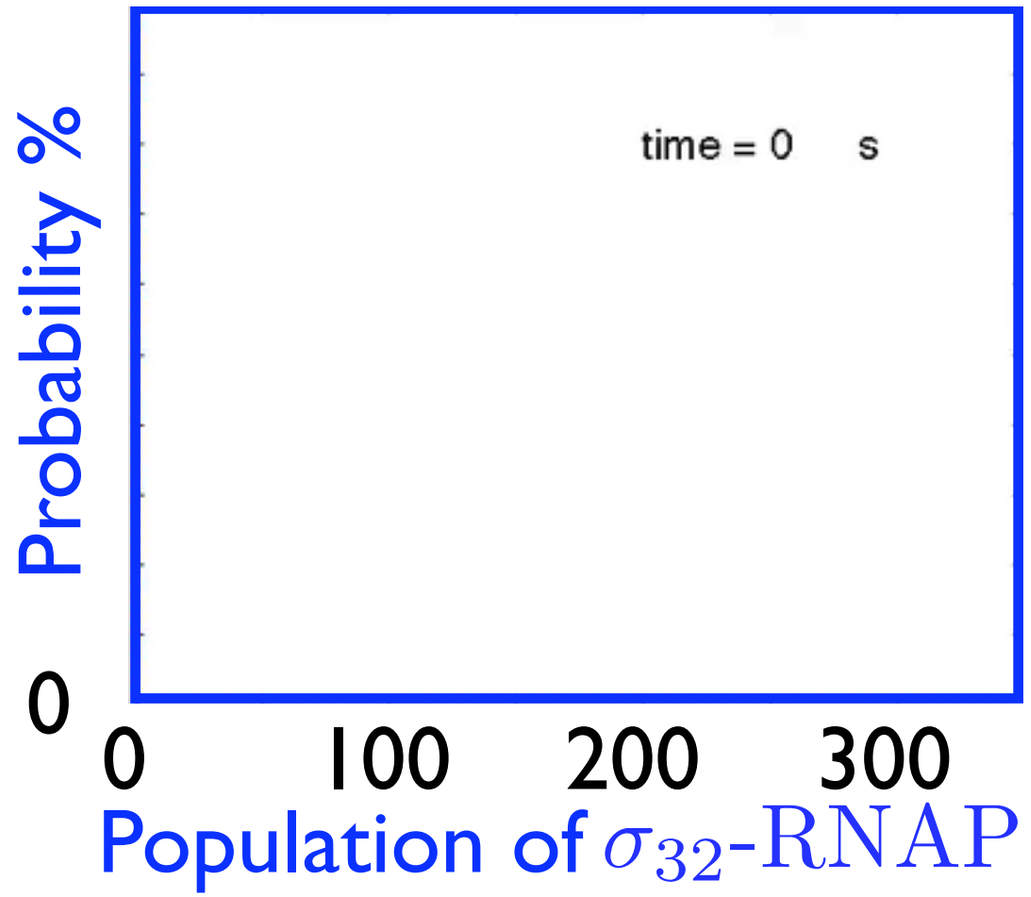
49 ODEs

165

Toy Heat Shock Model in *E. coli* (cont.)

Five Different FSP Solution Schemes:

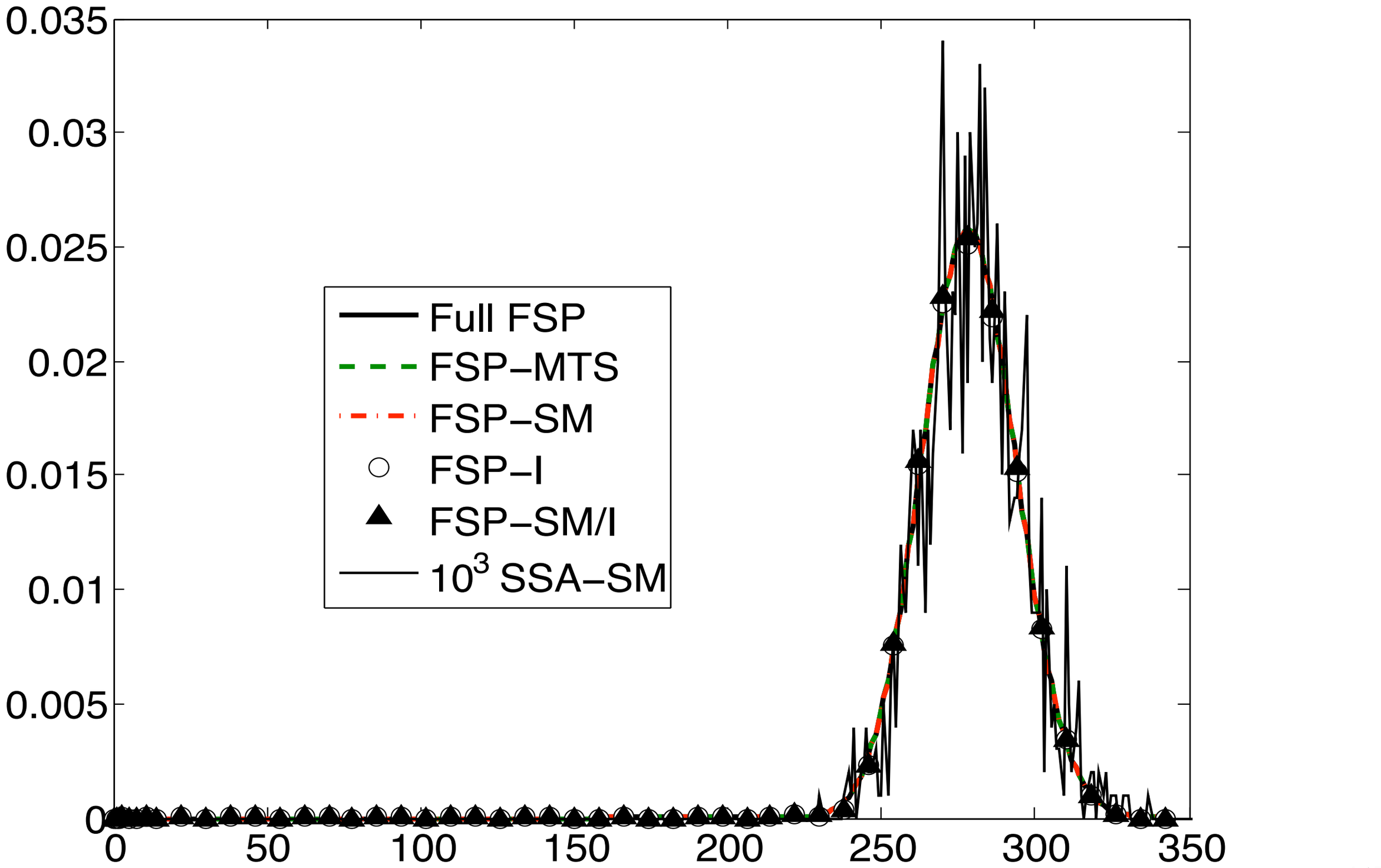
1. Full FSP
2. Slow manifold (FSP-SM)
3. Interpolated (FSP-I)
4. Hybrid (FSP-SM/I)
5. Multiple time interval (FSP-MTI)



70 sets of 195 or fewer ODEs.

166

Efficiency and accuracy of the reduced FSP methods



167

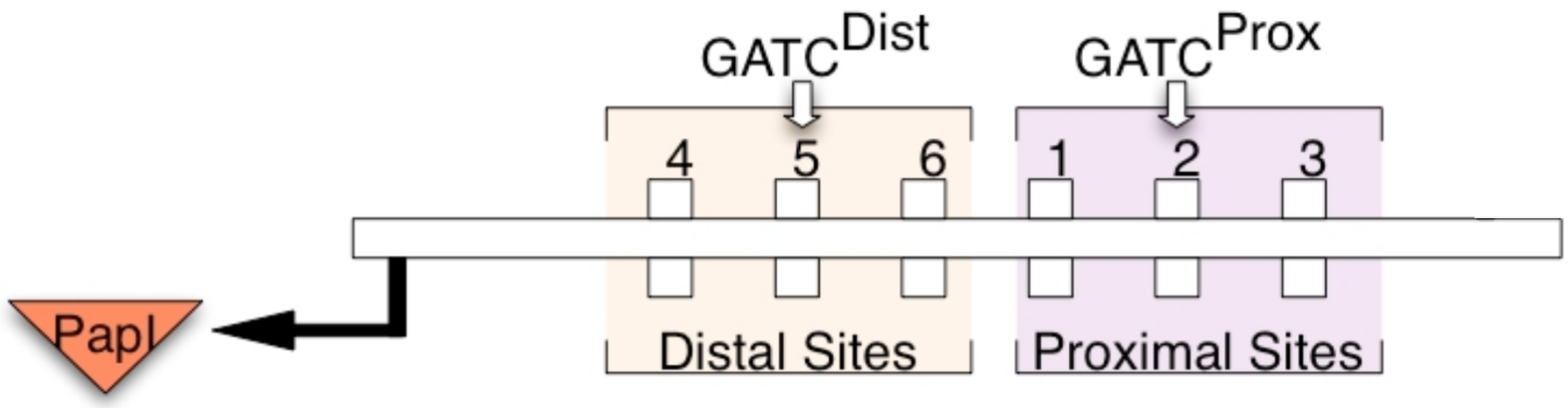
Efficiency and accuracy of the reduced FSP methods

For final time $t_f = 300s$				
Method	Matrix Size	J_{solve}	J_{total}	∞-norm Error
FSP	4459	750s	750s	$< 3.0 \times 10^{-5}$
FSP-MTS	195[1]	-	40.2s	$< 1.68 \times 10^{-4}$
FSP-SM	343	0.25s	0.94s	$\approx 5.1 \times 10^{-4}$
FSP-I	539	5.1s	6.1s	$\approx 7.7 \times 10^{-4}$
FSP-SM/I	49	0.04s	0.78s	$\approx 8.2 \times 10^{-4}$
10^4 SSA	Results would take more than 55 hours.			
10^3 SSA-SM	-	-	84.1s	≈ 0.0116
10^4 SSA-SM	-	-	925s	$\approx 3.4 \times 10^{-3}$
10^5 SSA-SM	-	-	9360s	$\approx 1.6 \times 10^{-3}$

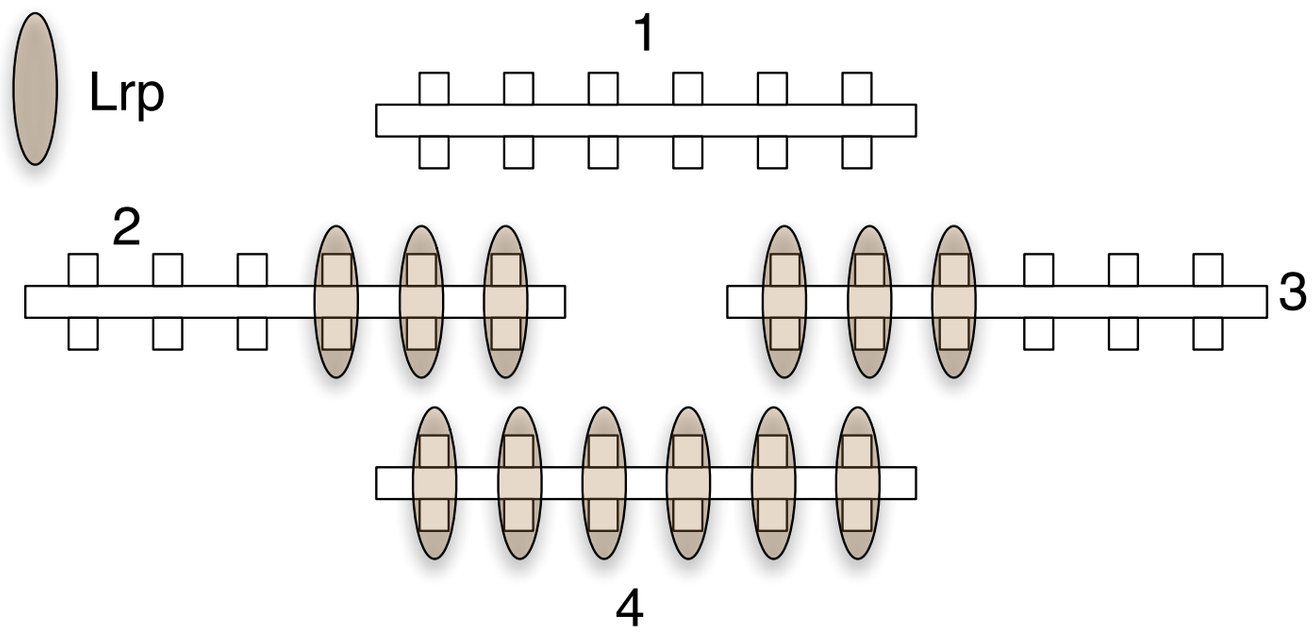
The Reduced FSP approaches are much faster and more accurate than alternative approaches!

168

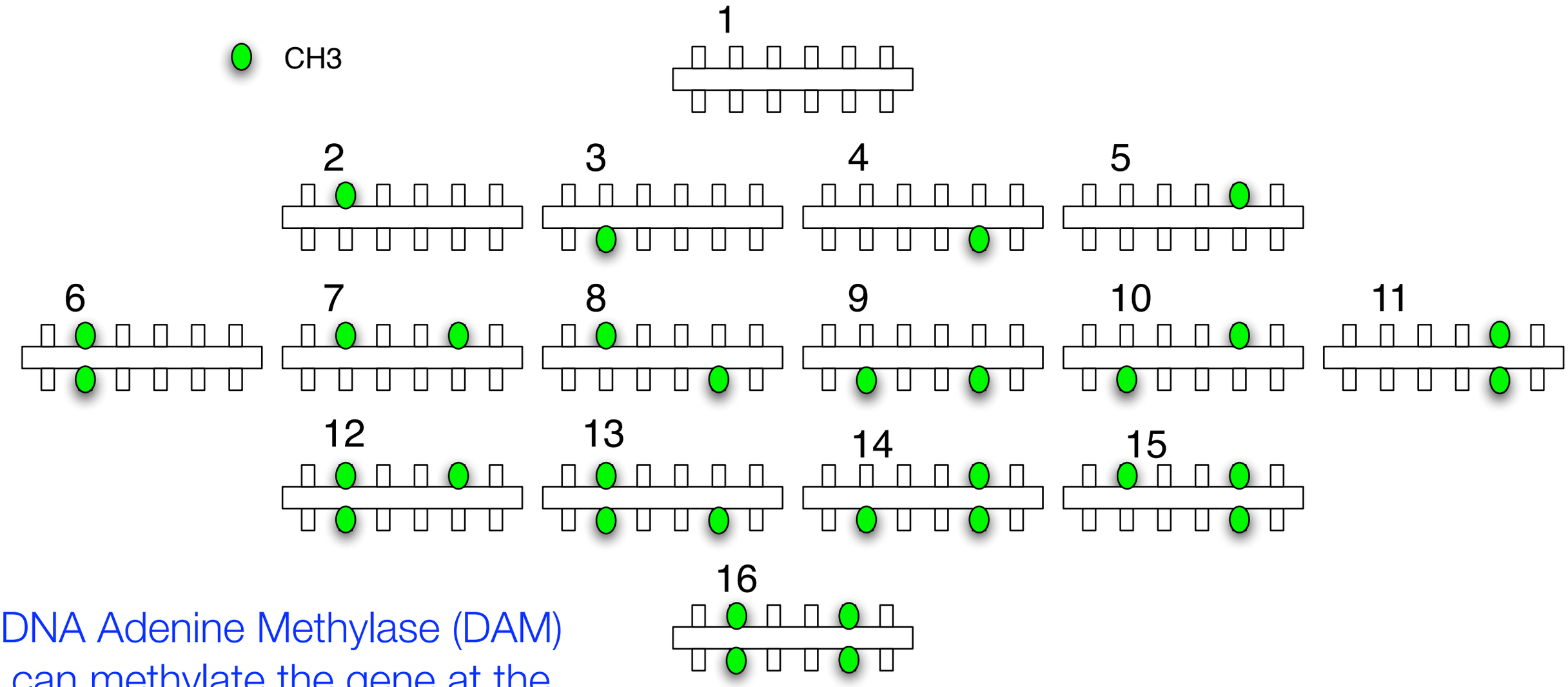
A More Realistic Pap Switch Model



4 gene states based on Lrp binding sites



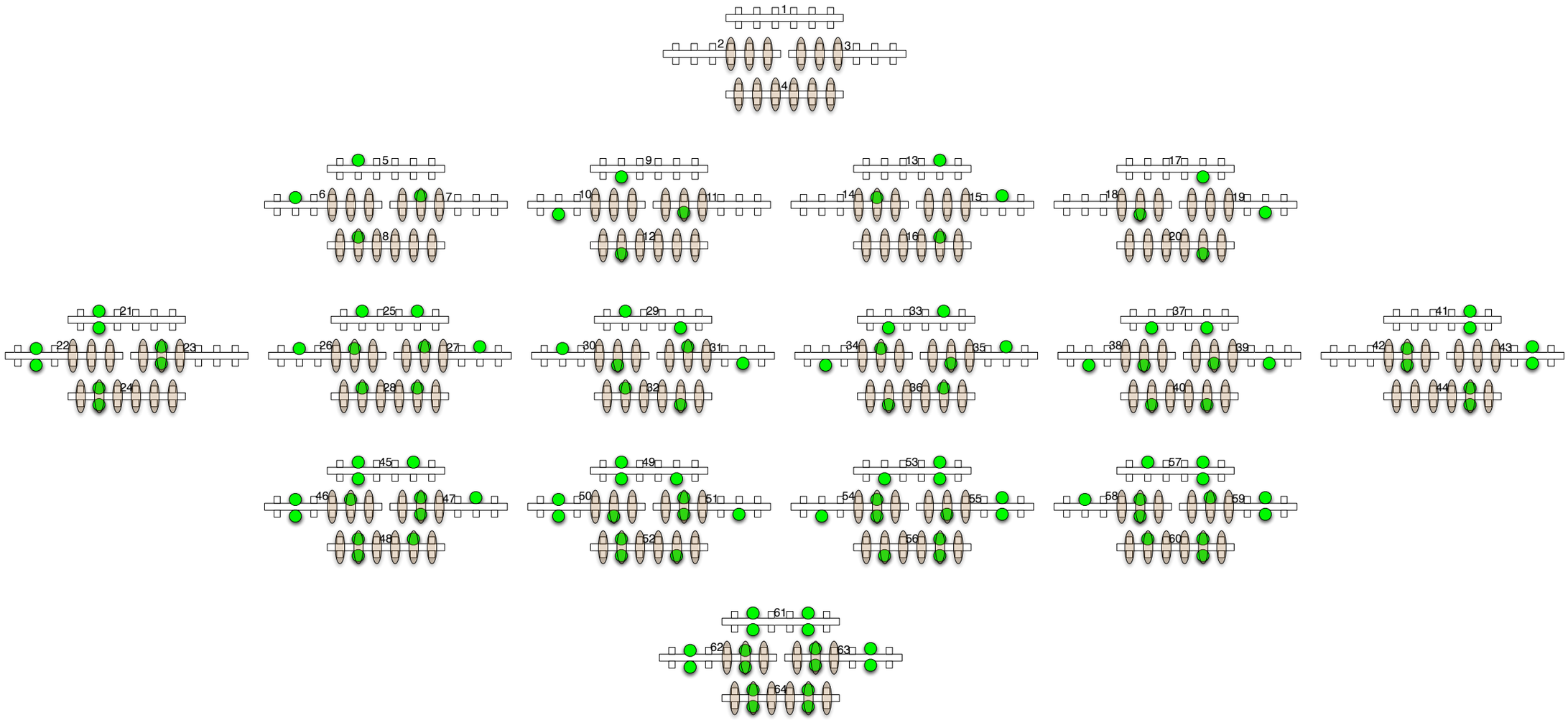
A More Realistic Pap Switch Model



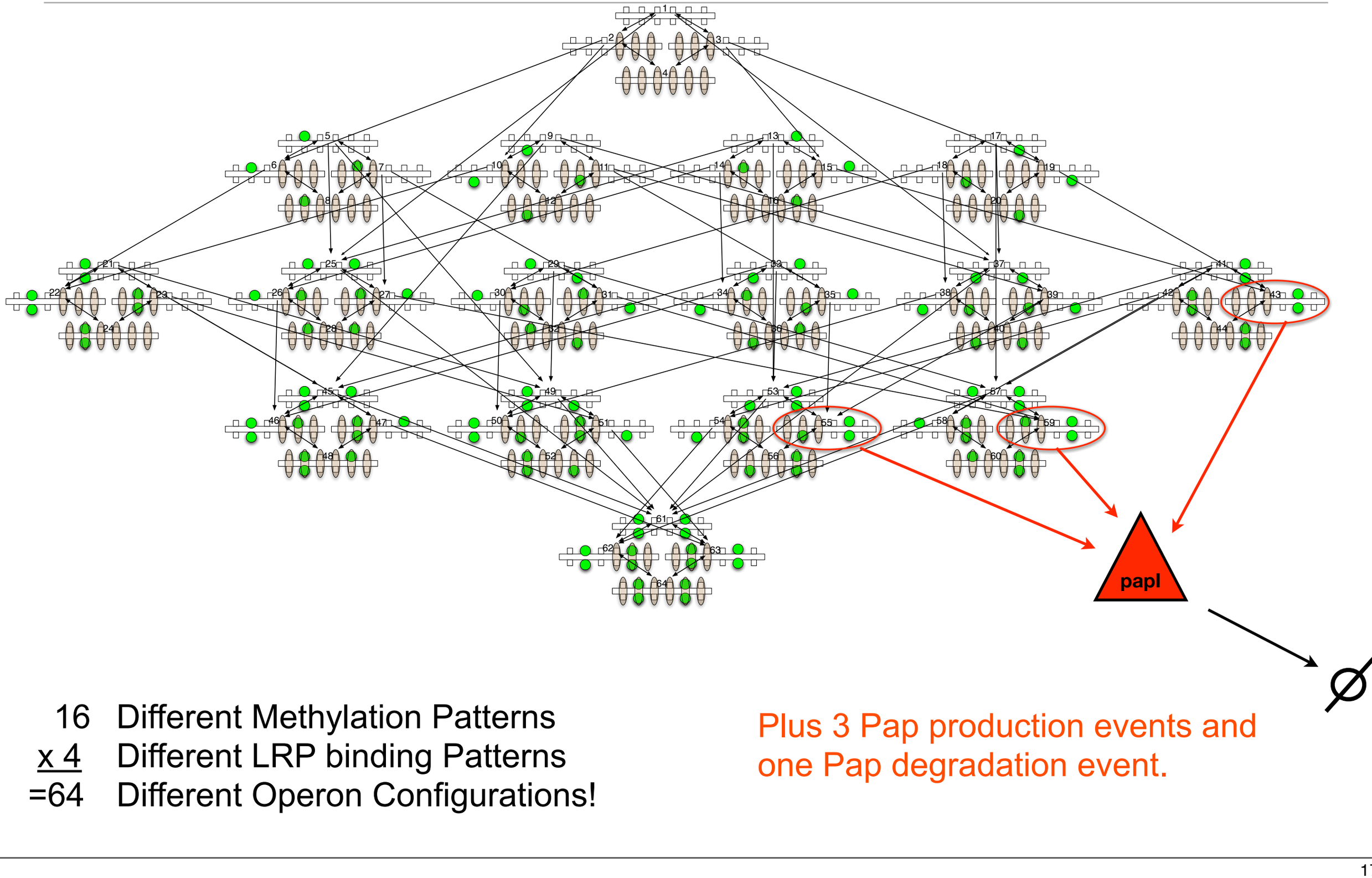
DNA Adenine Methylase (DAM) can methylate the gene at the GATC site

16 different possible methylation patterns

A More Realistic Pap Switch Model



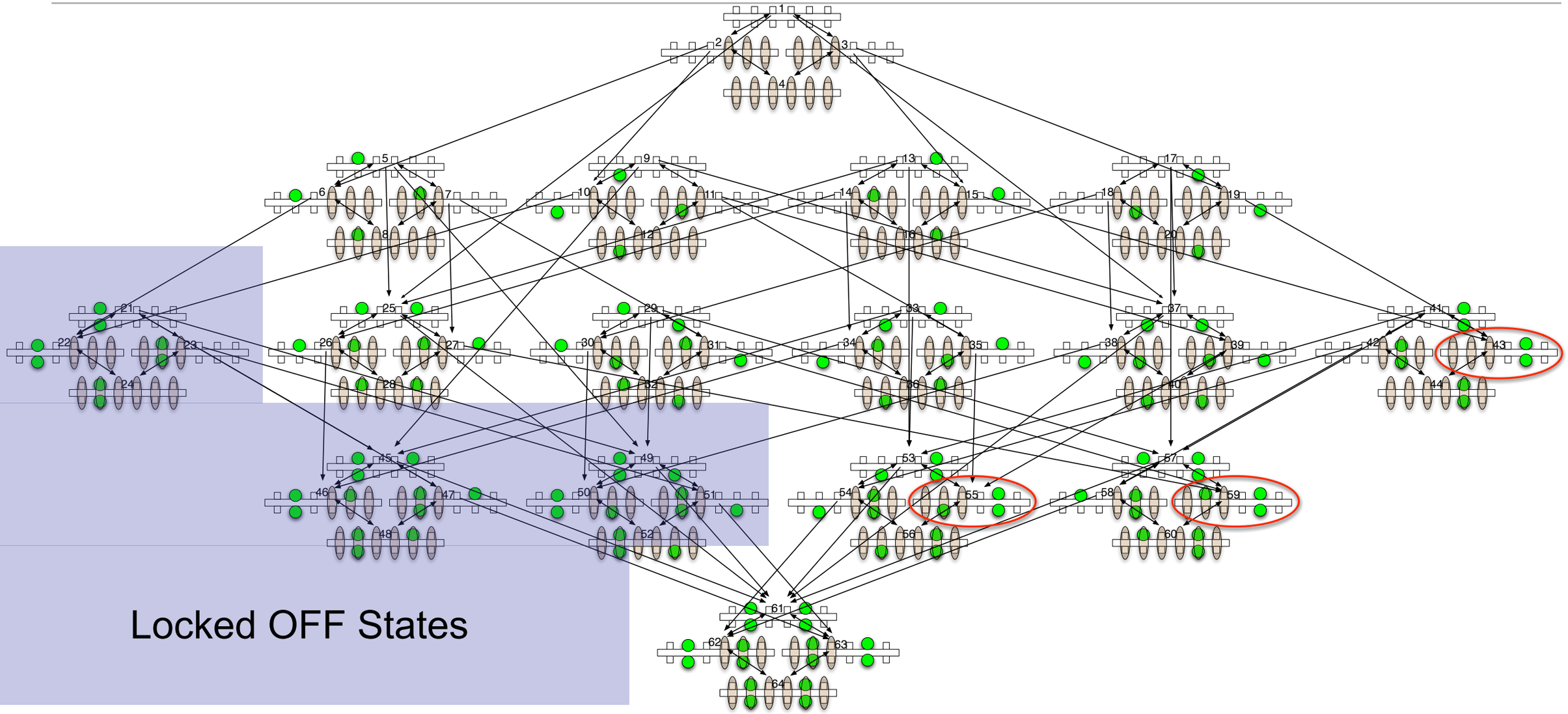
A More Realistic Pap Switch Model



16 Different Methylation Patterns
x 4 Different LRP binding Patterns
=64 Different Operon Configurations!

Plus 3 Pap production events and one Pap degradation event.

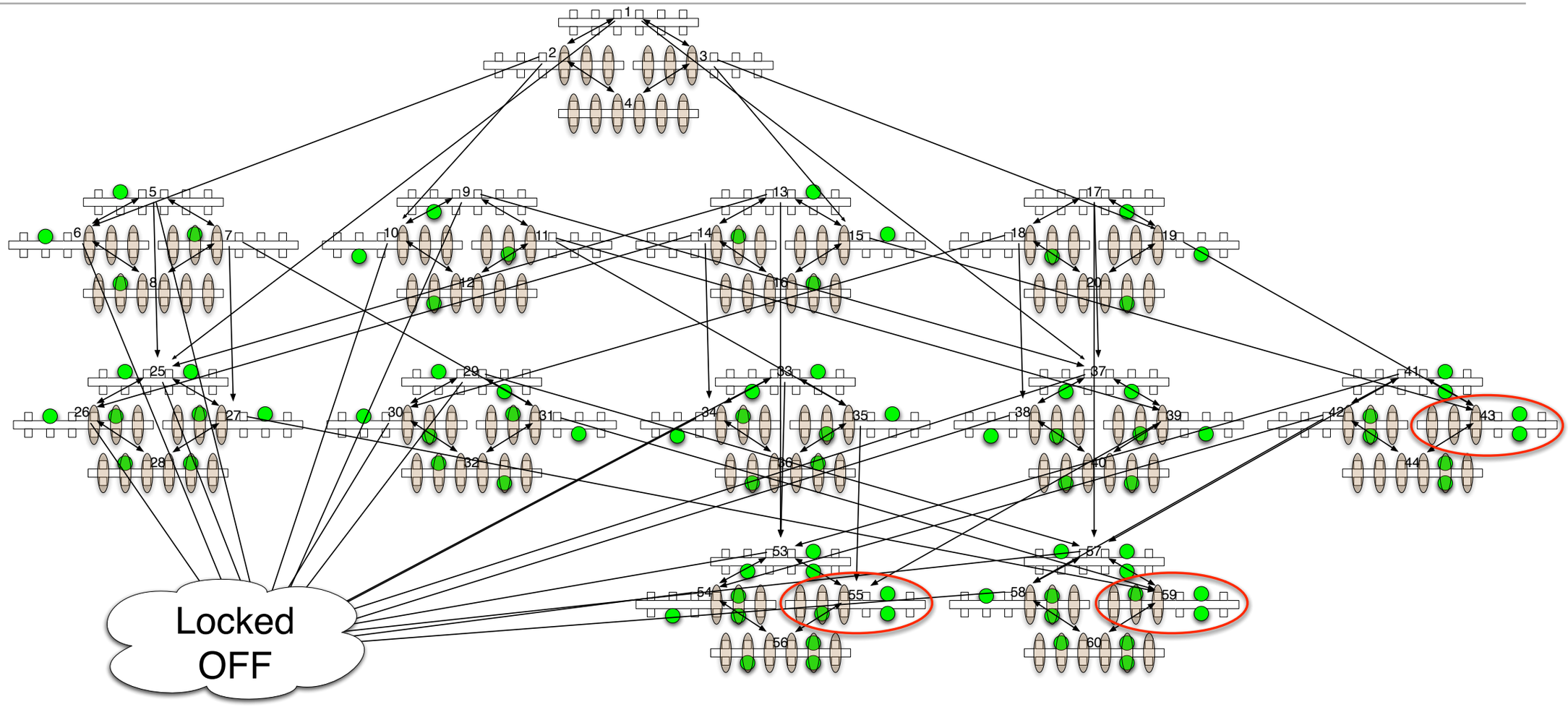
Aggregating Unobservable States



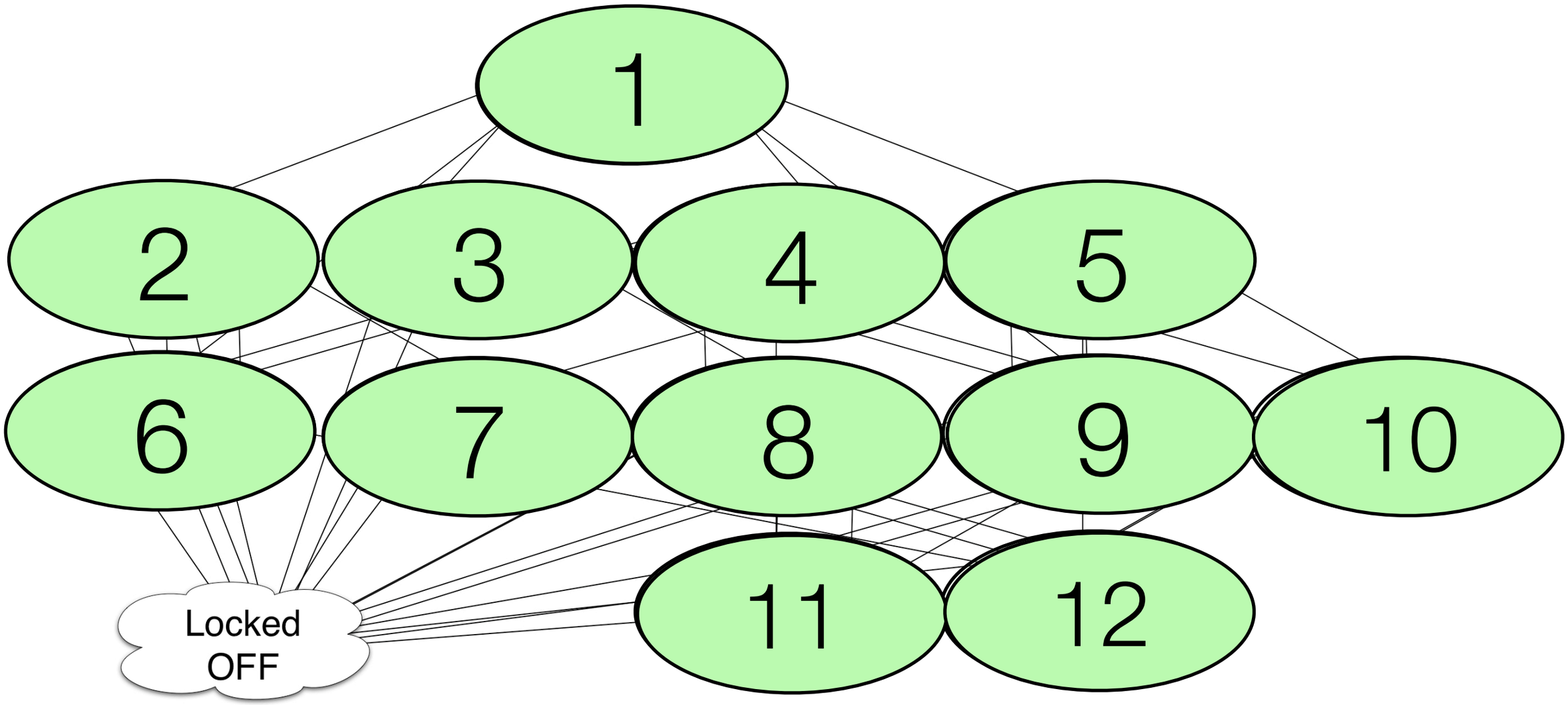
These States are unobservable from the ON states and can be aggregated.

16 Different Methylation Patterns
x 4 Different LRP binding Patterns
=64 Different Operon Configurations!

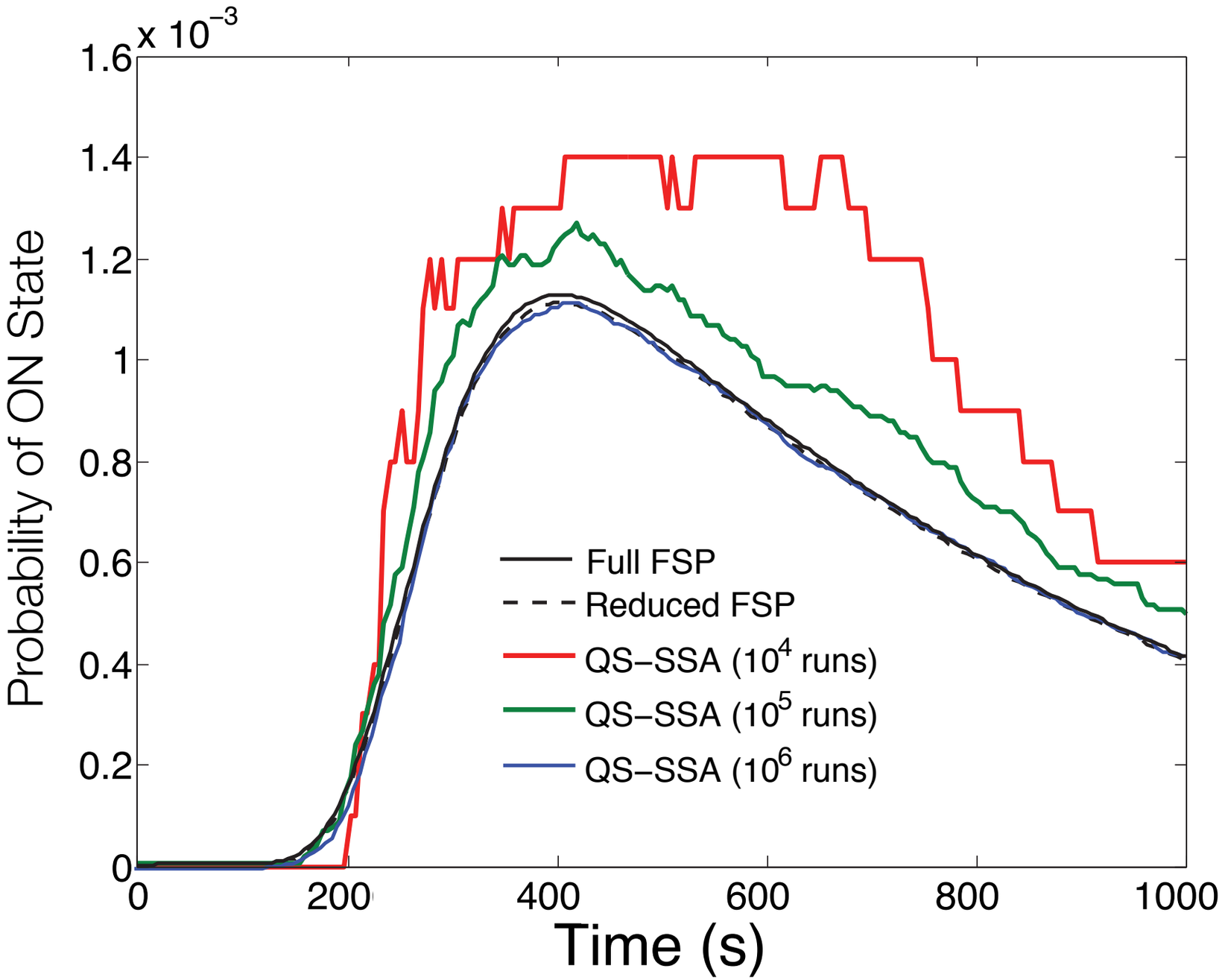
Aggregating Unobservable States



Aggregating Fast States



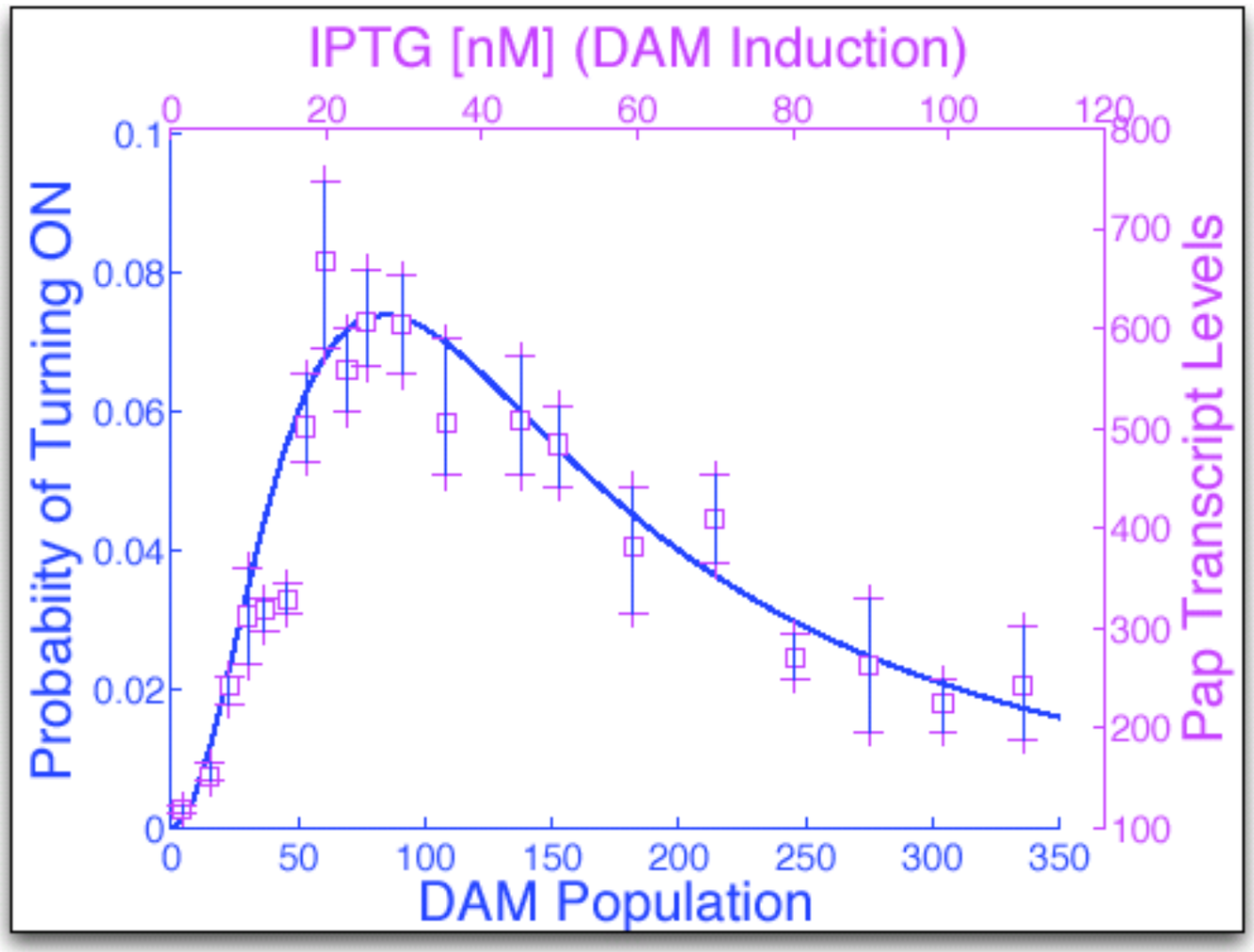
Comparisons



FSP vs. Monte Carlo Algorithms

Method	# Simulations	Time (s)[a]	Relative Error[b]
Full Model			
FSP	N.A. [c]	42.1	$< 0.013\%$
SSA	10^4	> 150 days	Not available
Reduced Model			
FSP approx.	N.A.	3.3	$\approx 1.3\%$
SSA approx.	10^4	9.8	$\approx 16\%$
SSA approx.	10^5	94.9	$\approx 7.7\%$
SSA approx.	10^6	946.2	$\approx 1.6\%$

Prediction vs. Experiments



Blue Line -- Predicted probability of expressing Pap versus DAM population.

Magenta -- Experimentally measured Pap transcript levels under different DAM induction levels.

Prediction vs. Experiments

Gene Alterations	Wild Type (100 Molecules)		Low Dam (25 Molecules)		High DAM (400 Molecules)	
	Exp.	Predicted (rate)	*Exp.*	Predicted (rate)	*Exp.*	Predicted (rate)
Wild Type	***Switching***	Switching (**7.2%**)	***OFF***	OFF (**2.6%**)	***OFF***	OFF (**1.2%**)
(1) ¼ proximal LRP affinity	***ON***	ON (**35.5 %**)	***ON***	ON (**22.2%**)	***ON***	ON (**12.5 %**)
(2) ¼ distal LRP affinity	***OFF***	OFF (**0.1%**)	***OFF***	OFF (**1.1%**)	***OFF***	OFF (**$4x10^{-4}$ %**)
(3) $GCTC^{prox}$	***OFF***	OFF (**0.0%**)	***OFF***	OFF (**0.0%**)	***OFF***	OFF (**0.0%**)
(4) $GCTC^{dist}$	***ON***	ON (**26.0%**)	***OFF***	OFF (**2.7%**)	***ON+***	ON+ (**88.7%**)

Effect of DAM concentration on the switching behavior of the wild type *pap* operon and four mutants. For every case, experimental observations (Hernday et al., '02) and model predictions are in agreement.

Conclusions

- Stochastic fluctuations or “noise” is present in the cell
  - Random nature of reactions
  - Quantization of reactants
  - Low copy numbers
- Fluctuations may be very important
  - Cell variability
  - Cell fate decisions
- Some tools are available
  - Monte Carlo simulations (SSA and variants)
  - Moment approximation methods
  - Linear noise approximation (Van Kampen)
  - Finite State Projection
- Many more are needed!

Conclusions

The Finite State Projection: a new tool for stochastic analysis of gene networks

Advantages:

- Accuracy: solutions with a guaranteed error bounds
  Particularly suitable for studying rare events
- Speed: solutions can be much faster than Monte Carlo methods
  especially when the system has large number of reactions/reaction firings
- Insight: Provides valuable information at little additional cost:
  Sensitivity/robustness to parameter changes
  Effect of changes in initial probabilities

Limitations

- Scalability: Not feasible when there are many species with broad distributions (over the time of interest [0, t])